

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Méthodes de faisceaux en programmation non convexe

Desaegher, Fabrice

Award date:
2004

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



FUNDP
Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B-5000 Namur Belgique

Méthodes de faisceaux en Programmation non convexe

Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
par

Fabrice DESAEGHER

Promoteur : Prof. J.-J. Strodjot

Année Académique 2003-2004

Résumé

Le but de ce travail est de présenter des méthodes de faisceaux pour résoudre des problèmes de programmation convexe et non convexe. A cette fin, nous commençons par quelques rappels indispensables en optimisation non différentiable, puis nous présentons deux méthodes de faisceau en optimisation convexe : une méthode de type primale et une méthode basée sur le concept de trajectoire proximale. Dans une seconde partie, nous abordons le cas non convexe en étudiant deux autres méthodes de faisceaux, l'une utilisant le concept de trajectoire proximale, l'autre utilisant la différence de deux fonctions convexes. Finalement nous présentons des résultats numériques pour illustrer le comportement de l'algorithme dans le cas non convexe.

Abstract

The aim of this work is to present bundle methods for solving convex and nonconvex programming problems. To this end, we begin by some indispensable recalls in nonsmooth optimization, then we present two bundle methods in convex optimization : a primal bundle method and a method based on the proximal trajectory. In a second part, we consider the non convex case by presenting two other bundle methods, one using the proximal trajectory and the other using the difference between two convex functions. Finally we report some numerical results to illustrate the behavior of the algorithm in the nonconvex case.

Table des matières

Introduction	3
1 Fonctions convexes et Sous-différentiels	5
1.1 Définitions et Notations	5
1.2 Dérivées directionnelles et Sous-différentiels	7
1.3 Le Sous-différentiel comme Multifonction	10
1.4 Sous-différentiel et Limite de Gradients	11
1.5 Sous-différentiels approximés	11
1.6 Le sous-différentiel approximé comme Multifonction	13
1.7 Le cas non convexe	14
2 Méthodes Faisceaux	17
2.1 Approche primale des Méthodes Faisceaux	17
3 L'Algorithme de Trajectoire proximale pour la minimisation convexe	24
3.1 Introduction	24
3.2 Méthodes faisceaux et Trajectoire proximale	25
3.3 La trajectoire proximale et la fonction du plan sécant	28
3.4 La méthode	37
4 Minimisation de fonctions non différentiables via des plans sécants et le controle de proximité	45
4.1 Introduction	45
4.2 Le modèle	46
4.3 L'algorithme	52
4.4 La convergence	55
4.5 Implémentation pratique	60

5	Une technique faisceau en minimisation non convexe non différentiable	62
5.1	Introduction	62
5.2	Le modèle	63
5.3	L'algorithme	68
5.4	La convergence	72
5.5	Conclusions	77
6	Résultats Numériques	80
6.1	Problèmes tests	80
6.2	Résultats numériques	83

Introduction

Dans les applications pratiques en optimisation, on se trouve souvent dans la situation où la fonction objectif à minimiser ou à maximiser n'est pas nécessairement différentiable. Un exemple typique est le problème de minimisation d'une fonction pouvant s'exprimer comme le maximum d'un nombre fini (ou infini) de fonctions (non) différentiables :

$$(P) \begin{cases} \min & f(x) = \max_{i \in I} f_i(x) \\ \text{s.c.} & x \in C. \end{cases}$$

Dans ce mémoire, nous ne considérerons que des problèmes dans des espaces de dimensions finies. Donc, pour chaque $i \in I$, les fonctions f_i sont définies sur \mathbb{R}^n et l'ensemble C est un sous-ensemble non vide de \mathbb{R}^n . Même si chaque fonction f_i est différentiable, la fonction f est, en général, non différentiable à chaque intersection entre les f_i .

Remarquons cependant que le problème non différentiable (P) peut être transformé en un problème différentiable en ajoutant une nouvelle variable v :

$$(\bar{P}) \begin{cases} \min & v \\ \text{s.c.} & v \geq f_i(x), \quad i \in I \\ & x \in C. \end{cases}$$

Les problèmes (P) et (\bar{P}) sont équivalents dans le sens où

$$x^* \text{ est une solution de } (P) \implies (v = f(x^*), x^*) \text{ est une solution de } (\bar{P})$$

$$(v^*, x^*) \text{ est une solution de } (\bar{P}) \implies x^* \text{ est une solution de } (P) \text{ et } v^* = f(x^*).$$

Observons que le problème (P) est non différentiable, tandis que le problème (\bar{P}) est différentiable. Néanmoins, lorsque I est de dimension infinie, ou lorsque I est un très grand ensemble discret, le problème (\bar{P}) a un nombre infini de contraintes, ou un très grand nombre de contraintes. Ce type de problème (appelé problème de programmation semi-infinie) n'est pas facile à résoudre, même si les données sont différentiables. Il est donc utile d'élaborer des algorithmes capables de résoudre le problème (P) et qui prennent

directement en compte la non différentiabilité.

Le but de ce mémoire est d'utiliser les méthodes faisceaux afin de résoudre le problème sans contraintes suivant :

$$\begin{array}{ll} \min & f(x) \\ \text{s.c.} & x \in \mathbb{R}^n \end{array}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ n'est pas nécessairement différentiable.

Nous commencerons tout d'abord, dans le premier chapitre, par rappeler quelques propriétés importantes des fonctions convexes, nécessaires à la compréhension du reste du mémoire.

Ensuite, dans le deuxième chapitre, nous introduirons la technique des méthodes faisceaux en minimisation convexe non différentiable.

Dans le troisième chapitre, on présentera une méthode faisceau incluant le concept de trajectoire proximale en minimisation convexe.

Dans les quatrième et cinquième chapitres, on présentera des méthodes faisceaux pour la minimisation non convexe.

Ces chapitres sont basés sur les articles [9], [10], [11] Enfin, dans le sixième chapitre, on exposera les résultats numériques portant sur l'implémentation de l'algorithme présent dans le chapitre 5.

Chapitre 1

Fonctions convexes et Sous-différentiels

Avant de pouvoir parler de méthodes numériques, il est nécessaire de rappeler quelques propriétés importantes des fonctions convexes. Les preuves seront omises. Cependant, le lecteur, s'il le désire, pourra les consulter dans [12].

1.1 Définitions et Notations

Tous les vecteurs sont considérés comme des vecteurs colonnes. On notera par

- $x^T y$ ou $\langle x, y \rangle$, le produit scalaire habituel dans \mathbb{R}^n des vecteurs x et y ;
- $\|\cdot\|$, la norme Euclidienne dans \mathbb{R}^n ;
- $B(x; r)$, la boule ouverte de centre x et de rayon $r > 0$;
- $[x, y]$ et (x, y) , le segment respectivement fermé et ouvert joignant x et y .

Définition 1.1.1 Un ensemble $C \subseteq \mathbb{R}^n$ est convexe si $[x, y] \subseteq C, \forall x, y \in C$.

Une combinaison linéaire $\sum_{j=1}^k \alpha_j x_j$ est appelée combinaison convexe des

points $x_1, \dots, x_k \in \mathbb{R}^n$ si chaque $\alpha_j \geq 0$ et $\sum_{j=1}^k \alpha_j = 1$.

L'enveloppe convexe d'un ensemble $C \subseteq \mathbb{R}^n$, notée par $\text{conv } C$, est l'ensemble de toutes les combinaisons convexes des points de C . C'est le plus petit ensemble convexe contenant C .

Définition 1.1.2 Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe si

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y),$$

chaque fois que x et y sont dans \mathbb{R}^n et que $\alpha \in [0, 1]$.

Si on a l'inégalité stricte avec $x \neq y$ et $\alpha \in (0, 1)$, alors f est dite strictement convexe.

f est dite fortement convexe de module $c > 0$ si $f - \frac{c}{2} \|\cdot\|^2$ est convexe.

Une fonction f est localement Lipschitzienne de constante L en $x \in \mathbb{R}^n$ s'il existe $\delta > 0$ tel que

$$|f(y) - f(z)| \leq L\|y - z\| \quad \forall y, z \in B(x; \delta).$$

Une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est semicontinue inférieurement en x si pour toute suite x_i convergeant vers x , on a :

$$f(x) \leq \liminf_{i \rightarrow \infty} f(x_i).$$

Notez qu'une fonction localement Lipschitzienne en x est continue en x .

Une première classe importante de fonctions convexes est la classe des fonctions affines $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de la forme $f(x) = a^T x + b$ où $a \in \mathbb{R}^n$ et $b \in \mathbb{R}$ (quand $b = 0$, une telle fonction est dite linéaire).

Les fonctions affines jouent un rôle important en analyse convexe car on peut minorer n'importe quelle fonction convexe par une fonction affine.

Rappelons maintenant quelques propriétés de différentiabilité.

Définition 1.1.3 La dérivée directionnelle de f en x dans la direction $d \in \mathbb{R}^n$ est définie par

$$f'(x; d) = \lim_{t \searrow 0} \frac{f(x + td) - f(x)}{t}.$$

Quand f est différentiable en x , la dérivée directionnelle existe dans toutes les directions d et est une fonction linéaire de d . On a donc la relation

$$f'(x; d) = \nabla f(x)^T d.$$

1.2 Dérivées directionnelles et Sous-différentiels

Dans cette section, on généralise la dérivée classique à des fonctions convexes mais non nécessairement différentiables.

Théorème 1.2.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe. Alors $\forall x \in \mathbb{R}^n$, f est localement Lipschitzienne en x (et donc continue en x).

Définissons maintenant le sous-gradient et le sous-différentiel d'une fonction convexe.

Définition 1.2.1 Le sous-différentiel d'une fonction convexe $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en $x \in \mathbb{R}^n$ est l'ensemble

$$\partial f(x) = \{g \in \mathbb{R}^n \mid f(y) \geq f(x) + g^T(y - x) \quad \forall y \in \mathbb{R}^n\}.$$

Les éléments de $\partial f(x)$ sont appelés sous-gradients de f en x .

Interprétation géométrique

L'inégalité définissant le sous-gradient $g \in \partial f(x)$:

$$\forall y \in \mathbb{R}^n \quad f(y) \geq f(x) + g^T(y - x)$$

Théorème 1.2.5 Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe, alors $\forall y \in \mathbb{R}^n$, on a

$$f(y) = \max \{f(x) + g^T(y - x) \mid x \in \mathbb{R}^n, g \in \partial f(x)\}.$$

Remarque 1.2.1 Ce théorème est la généralisation d'un résultat connu pour le gradient où si f est convexe alors

$$\forall x, y \in \mathbb{R}^n \quad f(y) \geq f(x) + \nabla f(x)(y - x).$$

On présente les conditions d'optimalité dans la proposition suivante.

Proposition 1.2.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et soit $x \in \mathbb{R}^n$.

Les trois propositions suivantes sont équivalentes :

- (i) x minimise f sur \mathbb{R}^n ;
- (ii) $0 \in \partial f(x)$;
- (iii) $f'(x; d) \geq 0 \quad \forall d \in \mathbb{R}^n$.

Observons que quand f est différentiable, la condition d'optimalité devient $\nabla f(x) = 0$.

Finalement, on a le résultat suivant concernant l'existence et l'unicité d'un minimum.

Proposition 1.2.2 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et soit C un ensemble convexe non vide de \mathbb{R}^n . Alors

- (i) Si f est strictement convexe, alors f a au plus un minimum sur C ;
- (ii) Si f est fortement convexe, alors f a un minimum unique sur C .

veut dire que g est la pente d'une fonction affine qui minore f et qui passe par le point $(x, f(x))$.

Commençons tout d'abord par rappeler le théorème des accroissements finis qui nous servira plus tard.

Théorème 1.2.2 (Accroissements finis) Soit $x, y \in \mathbb{R}^n$, $x \neq y$ et soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction Lipschitzienne sur un ensemble ouvert $U \subset \mathbb{R}^n$ telle que le segment de droite $[x, y] \subset U$. Alors, il existe un point $u \in (x, y)$ tel que

$$f(y) - f(x) \in \partial f(u)^T(y - x).$$

Présentons maintenant les relations entre la dérivée directionnelle et le sous-différentiel d'une fonction convexe. En particulier, chacun de ces deux concepts peut être déduit de l'autre. C'est ce qu'explique le théorème suivant.

Théorème 1.2.3 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe. Alors $\forall x \in \mathbb{R}^n$, on a

- (i) $f'(x; d) = \max \{g^T d \mid g \in \partial f(x)\} \quad \forall d \in \mathbb{R}^n$,
- (ii) $\partial f(x) = \{g \in \mathbb{R}^n \mid f'(x; d) \geq g^T d \quad d \in \mathbb{R}^n\}$,
- (iii) $\partial f(x)$ est un ensemble convexe compact et non vide contenu dans $B(0; L)$ où L est la constante de Lipschitz de f en x .

Quand f est différentiable, le sous-différentiel se réduit au gradient.

Théorème 1.2.4 Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe et différentiable en $x \in \mathbb{R}^n$, alors

$$\partial f(x) = \{\nabla f(x)\}.$$

Le résultat suivant est le résultat principal de cette section. Il permet de représenter une fonction convexe au moyen des sous-gradients.

1.3 Le Sous-différentiel comme Multifonction

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et considérons la multifonction

$$\partial f : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto \partial f(x).$$

Cette multifonction satisfait les propriétés suivantes :

Proposition 1.3.1 *La multifonction $\partial f(x)$ est monotone dans le sens où $\forall x_1, x_2 \in \mathbb{R}^n$, $\forall g_1 \in \partial f(x_1)$ et $g_2 \in \partial f(x_2)$,*

$$(g_2 - g_1)^T (x_2 - x_1) \geq 0.$$

Proposition 1.3.2 *La multifonction ∂f est fermée dans le sens où*

$$\left. \begin{array}{l} x_k \rightarrow x \\ g_k \rightarrow g \\ g_k \in \partial f(x_k) \quad \forall k \end{array} \right\} \Rightarrow g \in \partial f(x).$$

Lemme 1.3.1 *Soit S un sous-ensemble compact de \mathbb{R}^n .*

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est localement Lipschitzienne en tout point de S , alors f est Lipschitzienne sur S .

Proposition 1.3.3 *La multifonction ∂f est localement bornée dans le sens où pour tout sous-ensemble borné B de \mathbb{R}^n ,*

$$\partial f(B) \equiv \cup_{b \in B} \partial f(b) \text{ est borné dans } \mathbb{R}^n.$$

1.4 Sous-différentiel et Limite de Gradients

Théorème 1.4.1 (Rademacher) Soient $U \subset \mathbb{R}^n$ un ensemble ouvert et $f : U \rightarrow \mathbb{R}$ une fonction Lipschitzienne sur U . Alors f est différentiable presque partout sur U .

Le théorème de Rademacher est utile pour observer que les points où une fonction convexe n'est pas différentiable constitue un ensemble de mesure de Lebesgue nulle. Soit Ω_f cet ensemble de points, on a alors

$$\partial f(y) = \{\nabla f(y)\} \iff y \notin \Omega_f.$$

Par le théorie de la mesure de Lebesgue, on sait aussi que si $x \in \Omega_f$, alors il existe une suite $\{x_k\}$ dans le complément de Ω_f qui converge vers x . Cette suite est bornée et donc, par la proposition 1.2.2, la suite $\nabla f(x_k)$ est aussi bornée. Alors $f(x_k)$ a au moins un point limite s qui appartient à $\partial f(x)$ car $\partial f(x)$ est fermé.

Dans la proposition suivante, ces points limites sont utilisés pour construire le sous-différentiel de f en x .

Proposition 1.4.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe. Alors, $\forall x \in \mathbb{R}^n$,

$$\partial f(x) = \text{conv} \{s \in \mathbb{R}^n \mid \exists \{x_k\} \subseteq \Omega_f^c \text{ avec } x_k \rightarrow x, \nabla f(x_k) \rightarrow s\}.$$

1.5 Sous-différentiels approximés

Dans cette section, on présente une extension du sous-différentiel ainsi que quelques-unes de ses propriétés. Mais avant cela, nous avons besoin d'étendre la notion de dérivée directionnelle.

Définition 1.5.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et soit $\epsilon > 0$. La ϵ -dérivée directionnelle de f en x dans la direction $d \in \mathbb{R}^n$ est définie par

$$f'_\epsilon(x; d) = \inf_{t>0} \frac{f(x + td) - f(x) + \epsilon}{t}.$$

Géométriquement, le quotient $\frac{f(x+td)-f(x)+\epsilon}{t}$ peut être interprété comme étant la pente de la droite joignant le point $(0, f(x)-\epsilon)$ au point $(t, f(x+td))$. Donc par définition, $f'(x; d)$ est la plus petite pente de toutes ces droites.

Étendons maintenant la notion de sous-gradient et de sous-différentiel d'une fonction convexe, comme nous venons de le faire pour la dérivée directionnelle.

Définition 1.5.2 Soit $\epsilon \geq 0$. Alors le ϵ -sous-différentiel d'une fonction convexe $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est l'ensemble

$$\partial_\epsilon f(x) = \{g \in \mathbb{R}^n \mid f(y) \geq f(x) + g^T(y - x) - \epsilon \quad \forall y \in \mathbb{R}^n\}.$$

Chaque élément $g \in \partial_\epsilon f(x)$ est appelé un ϵ -sous-gradient de f en x .

Les principales propriétés de l' ϵ -sous-différentiel sont résumées dans le théorème suivant.

Théorème 1.5.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe. Alors

(i) $\partial_0 f(x) = \partial f(x)$;

(ii) $f'_\epsilon(x; d) = \max\{g^T d \mid g \in \partial_\epsilon f(x)\} \quad \forall d \in \mathbb{R}^n$;

(iii) $\partial_\epsilon f(x) = \{g \in \mathbb{R}^n \mid f'_\epsilon(x; d) \geq g^T d \quad \forall d \in \mathbb{R}^n\}$;

(iv) $\partial_\epsilon f(x)$ est un ensemble convexe, compact et non vide tel que $\|g\| \leq L \quad \forall g \in \partial_\epsilon f(x)$ où L est la constante de Lipschitz de f en x .

Remarque 1.5.1 Le sous-différentiel de f en x est un concept local tandis que l' ϵ -sous-différentiel de f en x avec $\epsilon > 0$ est un concept global dans le sens que si f est modifié loin de x , alors le sous-différentiel $\partial f(x)$ reste inchangé tandis que l' ϵ -sous-différentiel $\partial_\epsilon f(x)$, $\epsilon > 0$, peut être changé.

Finalement, l' ϵ -sous-différentiel peut être utilisé pour caractériser des solutions ϵ -optimales.

Proposition 1.5.1 Soit $\epsilon \geq 0$. On a la condition d'optimalité suivante :

$$0 \in \partial_\epsilon f(x) \iff \forall y \in \mathbb{R}^n \quad f(x) \leq f(y) + \epsilon.$$

1.6 Le sous-différentiel approximé comme Multifonction

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe et soit $\epsilon > 0$. Considérons la multifonction

$$\partial_\epsilon f : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto \partial_\epsilon f(x).$$

La multifonction $\partial_\epsilon f$ avec $\epsilon > 0$ a de plus belles propriétés que ∂f .

Proposition 1.6.1 $\partial_\epsilon f$ est fermé pour $\epsilon \geq 0$ dans le sens où

$$\left. \begin{array}{l} \epsilon_k \rightarrow \epsilon \\ x_k \rightarrow x \\ g_k \rightarrow g \\ g_k \in \partial_{\epsilon_k} f(x_k) \quad \forall k \end{array} \right\} \Rightarrow g \in \partial_\epsilon f(x).$$

Proposition 1.6.2 La multifonction $\partial_\epsilon f$ est localement bornée sur \mathbb{R}^n pour $\epsilon \geq 0$.

Soit $x \in \mathbb{R}^n$ et soit $\delta > 0$ et L la constante de Lipschitz de f sur $B(0; \delta)$. Alors

$$\forall \delta' < \delta, \forall y \in B(x; \delta'), \forall g \in \partial_\epsilon f(y) \quad \|g\| \leq L + \frac{\epsilon}{\delta - \delta'}.$$

La propriété suivante est valide seulement pour l' ϵ -sous-différentiel quand $\epsilon > 0$. Rappelons que cette propriété n'est pas satisfaite, en général, par le sous-différentiel.

Proposition 1.6.3 *La multifonction $\partial_\epsilon f$ est semicontinue inférieurement sur \mathbb{R}^n pour $\epsilon > 0$.
 Soit $\{x_k\}$ une suite dans \mathbb{R}^n convergeant vers $x \in \mathbb{R}^n$ et soit $g \in \partial_\epsilon f(x)$.
 Alors il existe une suite $\{g_k\}$ convergeant vers g et telle que $g_k \in \partial_\epsilon f(x_k) \quad \forall k$.*

1.7 Le cas non convexe

On peut généraliser les concepts différentiels afin de considérer aussi les fonctions nonconvexes Lipschitziennes.

Le fait que les fonctions soient non différentiables implique qu'elles ne possèdent pas de dérivées directionnelles classiques. Cela nous oblige donc à généraliser le concept de dérivée directionnelle. Après cela, les autres concepts se généralisent de façon analogue. On ne reprendra que ce dont nous aurons besoin par la suite. La généralisation peut être faite de plusieurs façons. Nous nous contenterons d'utiliser l'approche de **Clarke** (1983) dans le cas de la dimension finie.

Définition 1.7.1 (Clarke) *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction localement Lipschitzienne en un point $x \in \mathbb{R}^n$. La dérivée directionnelle généralisée de f en x dans la direction de $d \in \mathbb{R}^n$ est définie par*

$$f^\circ(x; d) = \limsup_{\substack{y \rightarrow x \\ t \searrow 0}} \frac{f(y + td) - f(y)}{t}.$$

Nous sommes maintenant prêts à généraliser le sous-différentiel pour des fonctions nonconvexes Lipschitziennes.

Notez que la définition est analogue à celle se trouvant dans le théorème 1.2.3(ii) pour les fonctions convexes, où la dérivée directionnelle est remplacée par la dérivée directionnelle généralisée.

Définition 1.7.2 (Clarke) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction localement Lipschitzienne en $x \in \mathbb{R}^n$.

Alors le sous-différentiel de f en x est l'ensemble

$$\partial f(x) := \{g \in \mathbb{R}^n \mid f^\circ(x; d) \geq g^T d \quad \forall d \in \mathbb{R}^n\}.$$

Chaque élément $g \in \partial f(x)$ est appelé un sous-gradient de f en x .

Définition 1.7.3 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction localement Lipschitzienne en $x \in \mathbb{R}^n$.

On dit que f est faiblement semi-smooth si

$$\forall x \in \mathbb{R}^n \quad \forall d \in \mathbb{R}^n \quad \left\{ \begin{array}{l} f^\circ(x; d) \text{ existe} \\ [t_j \searrow 0 \quad g_j \in \partial f(x + t_j d)] \Rightarrow \langle g_j, d \rangle \rightarrow f^\circ(x; d) \end{array} \right.$$

Voici maintenant quelques propriétés du sous-différentiel.

Théorème 1.7.1 Soit f une fonction Lipschitzienne en x de constante L . Alors

(i) $\partial f(x)$ est un ensemble convexe, compact, non vide tel que $\partial f(x) \subset B(0; L)$;

(ii) $f^\circ(x; d) = \max\{g^T d \mid g \in \partial f(x)\} \quad \forall d \in \mathbb{R}^n$.

Lemme 1.7.1 Si f est différentiable en x , alors f est localement Lipschitzienne en x .

Théorème 1.7.2 Si f est différentiable en x , alors

$$\partial f(x) = \{\nabla f(x)\}.$$

Le théorème suivant montre que le sous-différentiel des fonctions Lipschitziennes est une généralisation du sous-différentiel des fonctions convexes.

Théorème 1.7.3 *Si la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est convexe, alors*

$$(i) \ f'(x; d) = f^o(x; d) \quad \forall d \in \mathbb{R}^n \text{ et}$$

$$(ii) \ \partial_c f(x) = \partial f(x),$$

où $\partial_c f(x)$ est le sous-différentiel de la fonction convexe f en x .

Chapitre 2

Méthodes Faisceaux

Dans ce chapitre, nous allons introduire une famille d'algorithmes, appelés méthodes faisceaux, pour minimiser une fonction convexe non différentiable. Selon que le sous-différentiel ou la fonction objectif est approximé, on obtiendra le point de vue dual ou primal de ces méthodes. Ces approximations seront construites sous l'hypothèse suivante :

Hypothèse de base. En tout point $y \in \mathbb{R}^n$, seule la valeur $f(y)$ et un sous-gradient $g(y) \in \partial f(y)$ sont connus (par le biais d'un "oracle").

Les preuves des méthodes présentées dans ce chapitre seront omises. De plus, nous ne verrons que le point de vue primal des méthodes faisceaux car nous n'aurons pas besoin du point de vue dual par la suite.

2.1 Approche primale des Méthodes Faisceaux

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe. On va utiliser les sous-gradients donnés par l'oracle afin de définir des fonctions affines minorantes. Supposons que pendant le processus d'optimisation, on ait obtenu des points $y^j \in \mathbb{R}^n$, $j = 1, \dots, k$ et les informations correspondantes $f(y^j)$ et $g^j \in \partial f(y^j)$ grâce à l'oracle. Supposons que x^k est le point courant de l'itération. Alors, considérons la fonction convexe définie pour $x \in \mathbb{R}^n$ par

$$\varphi^k(x) = \max\{f(y^j) + (g^j)^T(x - y^j) \mid j \in J_k\},$$

où $J_k \subseteq \{1, \dots, k\}$. La fonction φ^k minore f car, par l'inégalité du sous-gradient, $f(x) \geq f(y^j) + (g^j)^T(x - y^j)$ pour tout x . Cette fonction est appelée le *modèle du plan sécant*. En utilisant les erreurs de linéarisation

$$\alpha_j^k = f(x^k) - f(y^j) - (g^j)^T(x^k - y^j), \quad j \in J_k,$$

ce modèle peut être exprimé sous la forme

$$\varphi^k(x) = \max_{j \in J_k} \{f(x^k) + (g^j)^T(x - x^k) - \alpha_j^k\} \quad \forall x \in \mathbb{R}^n.$$

En utilisant ce modèle, on peut obtenir une direction de descente d^k en résolvant le problème suivant

↳ à quoi?

$$(CP) \quad d^k = \arg \min_d \{\varphi^k(x^k + d) - f(x^k)\}.$$

En ajoutant une nouvelle variable, ce problème peut être exprimé comme un problème de programmation linéaire ordinaire :

$$(CP') \quad \begin{cases} \min_{v,d} & v \\ \text{s.c.} & -\alpha_j^k + (g^j)^T d \leq v, \quad j \in J_k. \end{cases}$$

Une fois la direction d^k obtenue, on pose $y^{k+1} = x^k + d^k$ et on met à jour le modèle φ^{k+1} .

L'ensemble $\{(g^j, \alpha_j^k)\}_{j \in J_k}$ est appelé faisceau. Il représente une collection de sous-gradients (fournis par l'oracle) au point courant x^k .

La difficulté avec le modèle du plan sécant est que la solution du problème (CP) peut ne pas être bornée. Une stratégie pour éviter ce problème est d'ajouter une contrainte de norme sur d dans le problème (CP') . Le nouveau problème est

$$BT(x^k, \rho) \quad \begin{cases} \min_{v,d} & v \\ \text{s.c.} & -\alpha_j^k + (g^j)^T d \leq v, \quad j \in J_k \\ & \frac{1}{2} \|d\|^2 \leq \rho, \end{cases}$$

où $\rho > 0$ est le rayon de la région de confiance. Cependant, afin d'éviter des contraintes quadratiques, on ajoute un terme quadratique à la fonction objectif du problème (CP') . On obtient le nouveau problème

$$QP(x^k, \gamma) \quad \begin{cases} \min_{v,d} & v + \frac{1}{2} \gamma \|d\|^2 \\ \text{s.c.} & -\alpha_j^k + (g^j)^T d \leq v, \quad j \in J_k, \end{cases}$$

où $\gamma > 0$. Les problèmes $BT(x^k, \rho)$ et $QP(x^k, \gamma)$ sont équivalents dans le sens où si (v, d) est une solution de $BT(x^k, \rho)$ et γ est un multiplicateur associé à la contrainte quadratique, alors (v, d) est une solution de $QP(x^k, \gamma)$ et, réciproquement, si (v, d) est une solution de $QP(x^k, \gamma)$, alors (v, d) est une solution de $BT(x^k, \rho)$ avec $\rho = (1/2) \|d\|^2$. De plus, un grand γ correspond à un petit ρ et un petit γ correspond à un grand ρ . Puisqu'il est difficile de choisir la valeur de ρ et qu'il est plus facile de travailler avec γ , les méthodes

faisceaux préfèrent utiliser le problème $QP(x^k, \gamma)$ pour calculer la direction de descente.

Afin de voir pourquoi il est intéressant d'utiliser le problème $QP(x^k, \gamma)$, examinons tout d'abord son Lagrangien dual.

Proposition 2.1.1 *Le Lagrangien dual du problème $QP(x^k, \gamma)$ est*

$$QD(x^k, \gamma) \begin{cases} \min & \frac{1}{2} \left\| \sum_{j \in J_k} \lambda_j g^j \right\|^2 + \gamma \sum_{j \in J_k} \lambda_j \alpha_j^k \\ \text{s.c.} & \sum_{j \in J_k} \lambda_j = 1, \lambda_j \geq 0, j \in J_k. \end{cases}$$

Si on note λ_j^k , $j \in J_k$ la solution de $QD(x^k, \gamma)$, alors la solution (v^k, d^k) de $QP(x^k, \gamma)$ est donnée par

$$v^k = -\frac{1}{\gamma} \|\bar{g}^k\|^2 - \bar{\alpha}^k \quad \text{et} \quad d^k = -\frac{1}{\gamma} \sum_{j \in J_k} \lambda_j^k g^j$$

$$\text{où } \bar{g}^k = \sum_{j \in J_k} \lambda_j^k g^j \text{ et } \bar{\alpha}^k = \sum_{j \in J_k} \lambda_j^k \alpha_j^k.$$

On déduit immédiatement du problème du problème $QP(x^k, \gamma)$ que la solution en v est $v^k = \varphi^k(x^k + d^k) - f(x^k)$, la réduction attendue en f quand on passe de x^k vers $x^k + d^k$. Si $v^k = 0$, alors x^k minimise f car alors $\bar{g}^k = 0$, $\bar{\alpha}^k = 0$ et $\bar{g}^k \in \partial_{\bar{\alpha}^k} f(x^k)$. Sinon, $v^k < 0$ et on s'attend à ce que d^k soit une direction de descente. Néanmoins, il peut arriver que ce ne soit pas le cas car nous travaillons avec φ^k , une approximation de f . La stratégie dans les méthodes faisceaux récentes est la suivante :

Poser $y^{k+1} = x^k + d^k$.

- Si $f(y^{k+1}) \leq f(x^k) + mv^k$ pour un certain $m \in]0, 1[$ donné, alors la réduction sur f est suffisante et on fait un pas sérieux, en posant $x^{k+1} = y^{k+1}$.
- Sinon, la réduction n'est pas suffisante et on fait un pas nul, en mettant $x^{k+1} = x^k$. Dans ce cas, le nouveau modèle du plan sécant φ^{k+1} doit inclure le sous-gradient en y^{k+1} , i.e. $k+1 \in J_{k+1}$, afin d'obtenir une meilleure direction de descente.

La direction de descente est très facile à implémenter parce qu'elle se

réduit soit à faire tout un pas, soit à ne pas faire de pas. D'un autre côté, l'entièreté de la procédure peut être vue comme deux itérations emboîtées, où la première itération est définie par des pas nuls consécutifs cherchant une direction de descente convenable, et la dernière itération est définie par des pas sérieux qui produisent la descente de la fonction objectif. L'algorithme est résumé ci-dessous.

La Méthode Faisceau :

Soit $\gamma > 0$, $m \in]0, 1[$, $x^0 \in \mathbb{R}^n$.

1. Obtenir $f(x^0)$ et $g^0 \in \partial f(x^0)$ de la part de l'oracle.
2. Initialiser $J_0 = \{0\}$, $\alpha_0^0 = 0$.
3. Pour $k = 0, 1, \dots$
résoudre le problème $QP(x^k, \gamma)$ ou son dual $QD(x^k, \gamma)$ pour avoir v^k et d^k . *dir. de desc. pour f^k .*
4. Si $v^k = 0$ alors STOP, sinon
5. mettre $y^{k+1} = x^k + d^k$.
6. Obtenir $f(y^{k+1})$ et $g^{k+1} \in \partial f(y^{k+1})$ de la part de l'oracle.
7. Si $f(y^{k+1}) \leq f(x^k) + mv^k$, alors (pas sérieux)
 - (a) mettre $x^{k+1} = y^{k+1}$,
 - (b) mettre $\alpha_j^{k+1} = \alpha_j^k + f(x^{k+1}) - f(x^k) - (g^j)^T d^k \quad \forall j \in J_k$,
 - (c) mettre $\alpha_{k+1}^{k+1} = 0$.
8. Sinon (pas nul)
 - (a) mettre $x^{k+1} = x^k$,
 - (b) mettre $\alpha_{k+1}^{k+1} = f(x^k) - f(y^{k+1}) + (g^{k+1})^T d^k$.
9. Mettre $J_{k+1} \subseteq J_k \cup \{k+1\}$ avec $k+1 \in J_{k+1}$.

Afin d'obtenir un algorithme efficace, il reste à discuter du choix du paramètre γ et de la taille du faisceau. Faire varier γ convenablement est souvent essentiel pour une bonne performance de la méthode faisceau. Quand γ est grand, le terme $\frac{1}{2}\gamma\|d\|^2$ augmente rapidement lorsque $\|d\|$ croît. Ceci signifie que la solution du problème $QP(x^k, \gamma)$ sera plutôt petite. Au contraire, quand γ est petit, le terme $\frac{1}{2}\gamma\|d\|^2$ augmente lentement lorsque $\|d\|$ croît, et la solution de $QP(x^k, \gamma)$ sera plutôt grande.

Quelques techniques basées sur cette remarque ont été inventées. Quand trop de pas nuls sont effectués, γ devrait croître afin de forcer les points tests à

être plus près de x^k et de cette manière à obtenir un modèle encore plus précis. D'un autre côté, quand trop de petits pas sérieux sont effectués, γ devrait décroître pour permettre aux points tests de sonder plus loin de x^k et de faire de plus grands pas. De façon plus précise, la stratégie pourrait être la suivante :

Mise à jour de γ :

1. Quand on obtient un pas sérieux, γ est mis à jour si

$$f(y^{k+1}) - f(x^k) \leq \overline{m}v^k,$$

où $0 < m < \overline{m} < 1$, i.e. quand la réduction est très grande. Dans ce cas, γ_k est réduit (par exemple, $\gamma_{k+1} = \max\{\gamma_k/10, \gamma_{\min}\}$, où $\gamma_{\min} > 0$ est une borne inférieure de γ).

2. Quand on obtient un pas nul, γ est augmenté si

$$\alpha(x^k, y^{k+1}) \geq -10v^k,$$

i.e. quand y^{k+1} est loin de x^k . Dans ce cas, γ_k est augmenté (par exemple, $\gamma_{k+1} = 10\gamma_k$).

3. Finalement, quand plus de trois pas sont exécutés sans changement de valeurs de γ , alors $\gamma_{k+1} = \max\{\gamma_k/2, \gamma_{\min}\}$.

Dans la méthode faisceau que l'on vient de présenter, on ne fait rien pour empêcher la taille du faisceau, i.e. le nombre d'éléments dans J_k , de grandir indéfiniment. En outre, la méthode peut ne pas converger si on choisit J_k trop petit. Quand le faisceau contient trop d'éléments, la solution proposée par Kiwiel est d'enlever au moins deux éléments du faisceau (non associés avec le dernier pas sérieux), et d'ajouter une linéarisation agrégée au faisceau. De façon plus précise, soit $\{x^k\}$ et $\{y^k\}$ les suites engendrées par la méthode faisceau. Pour chaque k , soient $\{f(y^k)\}$ et $g^k \in \partial f(y^k)$ obtenus de la part de l'oracle, et soit φ^k le modèle en y^k , i.e.

$$\varphi^k(x) = \max\{l^j(x) \mid j \in J_k^c\} \quad \forall x \in \mathbb{R}^n,$$

où $l^j(x) = f(y^j) + (s^j)^T(x - y^j)$, $j \in J_k^c \subseteq \{0, \dots, k\}$. Alors on sait que

$$y^{k+1} = \arg \min\{\varphi^k(x) + \frac{\gamma_k}{2}\|x - x^k\|^2\}.$$

On note par B_k^c le faisceau courant en y^k , i.e. $B_k^c = \{l^j(x) \mid j \in J_k^c\}$, et par B_{\max} la taille maximum du faisceau. Supposons maintenant que $|J_k^c| \geq B_{\max}$.

Alors, on enlève au moins deux éléments du faisceau (non associés avec le dernier pas sérieux). On met à jour l'ensemble d'indices J_k^c et on considère la linéarisation agrégée

$$l_k^a(x) = \varphi^k(y^{k+1}) + \gamma_k(x^k - y^{k+1})^T(x - y^{k+1}).$$

On a immédiatement que

$$l_k^a(y^{k+1}) = \varphi^k(y^{k+1}) \quad \text{et} \quad \gamma_k(x^k - y^{k+1}) \in \varphi^k(y^{k+1}).$$

En fait, par définition de y^{k+1} , on a que $0 \in \partial l_k^a(x) = \partial \varphi^k(y^{k+1}) + \gamma_k(y^{k+1} - x^k)$, i.e. $\gamma_k(x^k - y^{k+1}) \in \varphi^k(y^{k+1})$. On introduit alors un faisceau agrégé $B_k^a = \{l_k^a(x)\}$ et un nouveau modèle associé à y^{k+1}

$$\varphi^{k+1}(x) = \max\{l_k^a(x), \max\{l^j(x) \mid j \in J_k^c\}\}.$$

De façon plus générale, on note par $B_k^a = \{l_k^a(x) \mid j \in J_k^a \subseteq \{0, \dots, k-1\}\}$. Avec ces notations, on peut exprimer la procédure d'agrégation comme suit :

Procédure d'Agrégation :

1. Choisir $B_{max} \geq 3$. Mettre $B_{-1}^c, B_0^a = \emptyset$ et $J_0^a = \emptyset$.
2. Choisir x^0 et calculer $f(x^0)$ et g^0 . Mettre $k = 0$.
3. Mettre à jour $B_k^c = B_{k-1}^c \cup \{l^k(x)\}$ et $J_k^c = \{0 \leq j \leq k \mid l^j(x) \in B_k^c\}$.
4. Choisir γ_k et résoudre le problème de programmation quadratique

$$Q(x^k, \gamma_k) \begin{cases} \min_{v, x} & v + \frac{1}{2} \gamma_k \|x - x^k\|^2 \\ \text{s.c.} & l^j(x) \leq v, \quad j \in J_k^c \\ & l_j^a(x) \leq v, \quad j \in J_k^a, \end{cases}$$

afin d'obtenir (v^{k+1}, y^{k+1}) et de vérifier si le pas est sérieux ou nul. Mettre $x^{k+1} = y^{k+1}$ dans le premier cas et $x^{k+1} = x^k$ dans le dernier cas.

5. Si $|B_k^c| + |B_k^a| < B_{max}$, alors mettre $B_{k+1}^a = B_k^a$ et aller en 6. Sinon, choisir $C_k \subseteq B_k^c \cup B_k^a$ tel que $|C_k| \geq 2$ et $l^j(x) \notin C_k$, où $j \in J_k^c$ est l'indice du dernier pas sérieux. Mettre $B_k^c = B_k^c \setminus C_k$, $B_{k+1}^a = (B_k^a \setminus C_k) \cup \{l_k^a(x)\}$, où

$$l_k^a(x) = \varphi^k(y^{k+1}) + \gamma_k(x^k - y^{k+1})^T(x - y^{k+1}),$$

et

$$J_{k+1}^a = \{0 \leq j \leq k \mid l_j^a(x) \in B_{k+1}^a\}.$$

6. Mettre $k = k + 1$ et aller en 3.

En théorie, il est possible de travailler avec un faisceau de deux éléments, i.e. avec les informations associées au dernier pas sérieux et au sous-gradient agrégé (prendre $B_{max} = 3$). Kiwiel a prouvé la convergence de la méthode faisceau dans ce cas. On espère bien sûr que des faisceaux plus riches donneront des résultats avec une convergence plus rapide, mais cette possibilité de limiter la taille du faisceau permet à l'utilisateur de gagner de la place mémoire et de la vitesse de convergence.

Chapitre 3

L'Algorithme de Trajectoire proximale pour la minimisation convexe

3.1 Introduction

Le problème de minimisation est le suivant :

$$\begin{cases} \min & f(x) \\ \text{s.c.} & x \in \mathbb{R}^n \end{cases} \quad (3.1)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction convexe à valeurs finies et non nécessairement différentiable, avec un ensemble X^* non vide de minima.

Ce chapitre est organisé comme suit :

Dans la section 2, on rappelle les idées de base des méthodes faisceaux et le concept de trajectoire proximale [2, 20].

Dans la section 3, on se concentre sur la trajectoire proximale d'une fonction du plan sécant.

Dans la section 4, on présente une nouvelle méthode faisceau.

On rappelle que le sous-différentiel d'une fonction convexe au point x est l'ensemble de sous-gradients de f en x , i.e. l'ensemble de vecteurs $g \in \mathbb{R}^n$ satisfaisant l'inégalité

$$f(y) \geq f(x) + g^T(y - x) \quad \forall y \in \mathbb{R}^n.$$

Tandis que le ϵ -sous-différentiel d'une fonction convexe f au point x est l'ensemble de ϵ -sous-gradients de f en x , i.e. l'ensemble de vecteurs $g \in \mathbb{R}^n$

satisfaisant l'inégalité

$$f(y) \geq f(x) + g^T(y - x) - \epsilon \quad \forall y \in \mathbb{R}^n.$$

3.2 Méthodes faisceaux et Trajectoire proximale

Des méthodes faisceaux ont été conçues pour être des versions stabilisées de la *méthode du plan sécant* proposée par Cheney et Goldstein [4] et Kelly dans [17].

A chaque itération, on maintient un “faisceau”, i.e. un ensemble de triplets $B = \{(x_j, f(x_j), g_j) \mid j \in J\}$, indicés par J , constitués par le point générique x_j , par la valeur de la fonction objectif $f(x_j)$ et par un sous-gradient arbitraire $g_j \in \partial f(x_j)$.

Selon B , la méthode du plan sécant calcule l'itéré suivant x^+ qui est le minimum de la *fonction du plan sécant*

$$\hat{f}(x) = \max_{j \in J} f_j(x) \quad (3.2)$$

obtenue en prenant le maximum point par point des $|J|$ linéarisations de f prises aux points $x_j, j \in J$

$$f_j(x) = f(x_j) + g_j^T(x - x_j), \quad j \in J.$$

Il est connu que cette approche n'est pas efficace et souffre d'une instabilité numérique.

Pour stabiliser la méthode du plan sécant, le centre de stabilité a été introduit : c'est un point près duquel on essaie de calculer l'itéré suivant x^+ . Ainsi, une version de base des méthodes faisceaux résout à chaque itération le programme quadratique suivant :

$$\begin{cases} \min & \alpha \hat{f}(x) + \frac{1}{2} \|x - y\|^2 \\ \text{s.c.} & x \in \mathbb{R}^n \end{cases} \quad (3.3)$$

où y est le centre de stabilité et $\alpha > 0$ est un paramètre de poids convenable.

Une approche équivalente consiste à projeter le centre de stabilité y sur un ensemble de niveau de la fonction \hat{f} , défini par le paramètre $\gamma < \hat{f}(y)$, en résolvant le problème suivant :

$$\begin{cases} \min & \frac{1}{2} \|x - y\|^2 \\ \text{s.c.} & \hat{f}(x) \leq \gamma \\ & x \in \mathbb{R}^n \end{cases} \quad (3.4)$$

Finalement une troisième possibilité est de minimiser la fonction \hat{f} dans la boule de rayon $\tau > 0$ autour de y , i.e.

$$\begin{cases} \min & \hat{f}(x) \\ \text{s.c.} & \|x - y\| \leq \tau \\ & x \in \mathbb{R}^n \end{cases} \quad (3.5)$$

On rappelle [5] que les trois paramètres α, γ, τ sont essentiels pour la convergence globale des méthodes faisceaux.

En observant les problèmes (3.3), (3.4) et (3.5), on note qu'ils ont tous pour but de trouver un compromis entre deux objectifs en conflit : on veut d'un côté, minimiser la fonction polyédrale \hat{f} et, de l'autre côté, minimiser la distance du nouvel itéré x^+ au centre de stabilité y . En d'autres mots, on a le problème non-linéaire sans contraintes à double objectif suivant :

$$\begin{cases} \min & \hat{f}(x) \\ & \|x - y\| \\ \text{s.c.} & x \in \mathbb{R}^n \end{cases} \quad (3.6)$$

Rappelons maintenant, en rapport avec le problème (3.6), la définition d'efficacité faible.

Définition 3.2.1 *Un point \tilde{x} est une solution faiblement efficace du problème (3.6) s'il n'existe aucun point $x \in \mathbb{R}^n$ tel que*

$$\hat{f}(x) < \hat{f}(\tilde{x}) \quad \text{et} \quad \|x - y\| < \|\tilde{x} - y\|.$$

Le théorème suivant est prouvé dans [2]. Il éclaire l'équivalence appréciable des trois approches décrites au-dessus.

Théorème 3.2.1 *Un point $\tilde{x} \in \mathbb{R}^n$ est une solution faiblement efficace au problème (3.6) ssi $\exists \alpha \geq 0$ [resp. $\gamma \leq \hat{f}(y)$] [resp. $\tau \geq 0$] tel que \tilde{x} est une solution optimale du problème (3.3) [resp. (3.4)] [resp. (3.5)].*

La construction de l'entièreté de l'ensemble des solutions faiblement efficaces du problème (3.6) revient à résoudre paramétriquement un des problèmes (3.3), (3.4) ou (3.5).

On se concentre ainsi sur le problème (3.3). On indique son minimiseur par $x(\alpha, \hat{f}, y)$ ou plus simplement par $x(\alpha)$ chaque fois qu'il n'y a aucune confusion possible.

$\forall \alpha \geq 0$, $x(\alpha)$ satisfait l'inégalité

$$\alpha \hat{f}(x(\alpha)) + \frac{1}{2} \|x(\alpha) - y\|^2 \leq \alpha \hat{f}(x) + \frac{1}{2} \|x - y\|^2 \quad \forall x \in \mathbb{R}^n$$

et la condition d'optimalité

$$0 \in \alpha \partial \hat{f}(x(\alpha)) + x(\alpha) - y, \quad (3.7)$$

qui peut être réécrite de la façon suivante :

$$x(\alpha) = y - \alpha \hat{g}(\alpha), \quad \text{pour un certain } \hat{g}(\alpha) \in \partial \hat{f}(x(\alpha)). \quad (3.8)$$

Définition 3.2.2 La fonction $\alpha \in [0, +\infty[\mapsto x(\alpha) \in \mathbb{R}^n$ est la trajectoire proximale de \hat{f} par rapport à y .

Bien sûr, n'importe quel point \tilde{x} est une solution faiblement efficace du problème (3.6) ssi $\exists \tilde{\alpha} \geq 0$ tel que $\tilde{x} = x(\tilde{\alpha})$.

Les propriétés de la trajectoire proximale sont entièrement décrites dans [20], où le centre de stabilité y est mis égal à zéro. Rappelons ici certaines d'entre elles :

1. pour un centre de stabilité y donné, la trajectoire proximale est une courbe continue sur tout le domaine $[0, +\infty[$;
2. la fonction $\alpha \in [0, +\infty[\mapsto \hat{f}(x(\alpha))$ est décroissante ;
3. $\lim_{\alpha \rightarrow \infty} \hat{f}(x(\alpha)) = \hat{f}^*$, où $\hat{f}^* \triangleq \inf_{x \in \mathbb{R}^n} \hat{f}(x) \in \mathbb{R} \cup \{-\infty\}$;
4. $\lim_{\alpha \rightarrow \infty} x(\alpha) = \hat{x}^*$, pour $\hat{x}^* \in \hat{X}^*$, à condition que l'ensemble $\hat{X}^* \triangleq \{\hat{x}^* \in \mathbb{R}^n \mid \hat{x}^* = \arg \min_{x \in \mathbb{R}^n} \hat{f}(x)\}$ soit non vide et $\hat{f}^* = \min_{x \in \mathbb{R}^n} \hat{f}(x)$.

Les résultats exposés jusqu'ici peuvent être facilement étendus à n'importe quelle fonction convexe (et non nécessairement polyédrale).

Pour résumer ces résultats, on observe que chaque point de la trajectoire proximale, en correspondance avec un centre de stabilité y , est une solution

faiblement efficace (3.2.1) du problème à double objectif (3.6) qui lui correspond. En outre, la trajectoire proximale est continue et définie de façon unique, elle émane du centre de stabilité et se termine en un point minimum (s'il en existe un).

3.3 La trajectoire proximale et la fonction du plan sécant

Etant donné un point $x \in \mathbb{R}^n$, il est connu que le sous-différentiel polyédral de \hat{f} en x est donné par l'enveloppe convexe des gradients des linéarisations actives en x , i.e.

$$\partial \hat{f}(x) = \text{conv} \{g_j \mid j \in I_x\}$$

où $I_x \triangleq \{j \in J \mid \hat{f}(x) = f_j(x)\}$ est l'ensemble *actif* de \hat{f} en x .

A partir de la définition de la fonction du plan sécant, et tenant compte de la convexité, on peut en déduire les propriétés suivantes. Voir [15] pour les preuves correspondantes.

Lemme 3.3.1 *Etant donnés deux points $v, w \in \mathbb{R}^n$, si $I_v \cap I_w \neq \emptyset$, alors pour $\theta \in]0, 1]$*

$$I_{x_\theta} = I_v \cap I_w,$$

où $x_\theta \triangleq \theta v + (1 - \theta)w$.

Théorème 3.3.1 *Etant donnés deux points $v, w \in \mathbb{R}^n$, si $I_v \cap I_w \neq \emptyset$, alors pour $\theta \in]0, 1]$*

$$\partial \hat{f}(x_\theta) = \partial \hat{f}(v) \cap \partial \hat{f}(w),$$

où x_θ est défini comme dans le lemme 3.3.1

Dans [20], il a été montré que la trajectoire proximale est linéaire par morceaux. Etablissons maintenant un algorithme permettant de la construire en résolvant une suite finie de programmes linéaires. La trajectoire proximale peut cependant être construite en utilisant des méthodes telles que celles

rapportées dans [14] et [20].

On prouve le théorème suivant :

Théorème 3.3.2 *Soient v et w deux points de la trajectoire proximale correspondant respectivement à $\alpha_v > 0$ et $\alpha_w > 0$, tels que $I_v \cap I_w \neq \emptyset$. Alors pour $\theta \in]0, 1]$, le point $x_\theta \triangleq \theta v + (1 - \theta)w$ appartient à la trajectoire proximale ssi*

$$\hat{g}(\alpha_v), \hat{g}(\alpha_w) \in \partial \hat{f}(v) \cap \partial \hat{f}(w), \quad (3.9)$$

où $\hat{g}(\alpha_v)$ et $\hat{g}(\alpha_w)$ sont les sous-gradients de \hat{f} respectivement en v et w satisfaisant la condition d'optimalité (3.8).

Preuve :

- \Leftarrow Montrons d'abord que la condition (3.9) garantit que x_θ appartient à la trajectoire, i.e. que l'entièreté du segment de droite, reliant v à w , appartient à la trajectoire.
On sait par hypothèse que

$$v = y - \alpha_v \hat{g}(\alpha_v) \quad \text{avec } \hat{g}(\alpha_v) \in \partial \hat{f}(v) \quad (3.10)$$

et

$$w = y - \alpha_w \hat{g}(\alpha_w) \quad \text{avec } \hat{g}(\alpha_w) \in \partial \hat{f}(w) \quad (3.11)$$

En multipliant (3.10) et (3.11) respectivement par $\theta > 0$ et $(1 - \theta) > 0$, puis en les sommant, on obtient :

$$\begin{aligned} \theta v + (1 - \theta)w &= \theta y + (1 - \theta)y - \theta \alpha_v \hat{g}(\alpha_v) - (1 - \theta) \alpha_w \hat{g}(\alpha_w) \\ \Leftrightarrow x_\theta - y &= -(\theta \alpha_v + (1 - \theta) \alpha_w) \hat{g}_\theta, \end{aligned}$$

où $\hat{g}_\theta \triangleq \frac{\theta \alpha_v}{\theta \alpha_v + (1 - \theta) \alpha_w} \hat{g}(\alpha_v) + \frac{(1 - \theta) \alpha_w}{\theta \alpha_v + (1 - \theta) \alpha_w} \hat{g}(\alpha_w)$. Puisque $\hat{g}(\alpha_v), \hat{g}(\alpha_w) \in \partial \hat{f}(v) \cap \partial \hat{f}(w)$ et que l'intersection de deux ensembles convexes est encore convexe, le vecteur \hat{g}_θ appartient à $\partial \hat{f}(v) \cap \partial \hat{f}(w)$ qui, par le théorème 3.3.1, est exactement $\partial \hat{f}(x_\theta)$. On en conclut donc que le point x_θ appartient à la trajectoire proximale.

- \Rightarrow En supposant que pour $\theta \in]0, 1]$ le point x_θ appartient à la trajectoire proximale, il est facile de vérifier que $x_\theta = x(\alpha_\theta)$, avec $\alpha_\theta = \theta \alpha_v + (1 - \theta) \alpha_w$ et $\hat{g}(\alpha_\theta) = \hat{g}_\theta \in \partial \hat{f}(x_\theta)$.
Par le théorème 3.3.1, $\partial \hat{f}(x_\theta) = \partial \hat{f}(v) \cap \partial \hat{f}(w)$, d'où $\hat{g}_\theta \in \partial \hat{f}(v) \cap$

$$\partial \hat{f}(w) \quad \forall \theta \in]0, 1].$$

De plus par définition de \hat{g}_θ ,

$$\lim_{\theta \rightarrow 1} \hat{g}_\theta = \hat{g}(\alpha_v)$$

et

$$\lim_{\theta \rightarrow 0} \hat{g}_\theta = \hat{g}(\alpha_w),$$

ce qui entraîne que les vecteurs $\hat{g}(\alpha_v)$ et $\hat{g}(\alpha_w)$ appartiennent à l'ensemble $\partial \hat{f}(v) \cap \partial \hat{f}(w)$ car cet ensemble est fermé (puisque le sous-différentiel est compact). \square

Ici le théorème 3.3.2 est utile afin de définir l'algorithme suivant permettant de construire la trajectoire proximale en partant d'un centre de stabilité donné (i.e; pour $\alpha = 0$).

On indique par

- $\tilde{\alpha}$, la valeur courante de α ,
- \tilde{x} , le point $x(\tilde{\alpha})$,
- \tilde{I} , l'ensemble actif correspondant, i.e. l'ensemble $\{j \in J \mid \hat{f}(\tilde{x}) = f_j(\tilde{x})\}$

Algorithme 3.3.1

1. Initialisation : $\tilde{\alpha} = 0$; $\tilde{x} = y$
2. Trouver la valeur maximale α^* de α telle que

$$\partial \hat{f}(x(\alpha)) \supseteq \partial \hat{f}(\tilde{x})$$

et

$$\hat{g}(\alpha) \in \partial \hat{f}(\tilde{x}).$$

3. Si $\alpha^* = \infty$, STOP. Sinon aller au pas 4.
4. Si $\partial \hat{f}(x(\alpha^*)) \supset \partial \hat{f}(\tilde{x})$, poser $\tilde{\alpha} := \alpha^*$, $\tilde{x} = x(\alpha^*)$ et aller au pas 2. Sinon poser $\tilde{\alpha} := \alpha^*$, $\tilde{x} = x(\alpha^*)$ et aller au pas 5.
5. Trouver la valeur maximale α^* de α telle que $\alpha^* > \tilde{\alpha}$,

$$\partial \hat{f}(x(\alpha)) \subset \partial \hat{f}(\tilde{x})$$

et

$$\hat{g}(\tilde{\alpha}) \in \partial \hat{f}(x(\alpha)).$$

6. Si $\alpha^* = \infty$, STOP. Sinon mettre $\tilde{\alpha} := \alpha^*$, $\tilde{x} = x(\alpha^*)$ et aller au pas 2.

Remarque 3.3.1 Le segment $[\tilde{x}, x(\alpha^*)]$, où le point $x(\alpha^*)$ est obtenu soit au pas 2, soit au pas 5, appartient à la trajectoire proximale par le théorème 3.3.2.

Remarque 3.3.2 Le pas 5 est de nature combinatoire car il peut réclamer l'exploration de tous les sous-ensembles propres de \tilde{I} .

Nous allons maintenant voir en détails comment implémenter les pas 2 et 5 de l'algorithme. Le pas 2 de l'algorithme exige la solution du problème

$$\begin{cases} \alpha^* = \max_{\alpha \geq 0} \alpha \\ \text{s.c.} \quad \hat{g}(\alpha) \in \partial \hat{f}(\tilde{x}) \\ \partial \hat{f}(x(\alpha)) \supseteq \partial \hat{f}(\tilde{x}) \end{cases} \quad (3.12)$$

qui peut être écrit sous la forme d'un problème de programmation mathématique de la façon suivante :

$$\begin{cases} \alpha^* = \max_{\alpha, \lambda, x, \hat{g}} \alpha \\ \text{s.c.} \quad \hat{f}(x) = f_j(x), j \in \tilde{I} \\ \hat{f} \geq f_j(x), j \notin \tilde{I} \\ x = -\alpha \hat{g} + y \\ \hat{g} = \sum_{i \in \tilde{I}} \lambda_i g_i \\ \sum_{i \in \tilde{I}} \lambda_i = 1 \\ \alpha \geq 0 \quad \lambda_i \geq 0, i \in \tilde{I}. \end{cases} \quad (3.13)$$

Soit $\nu = \hat{f}(x)$, le problème (3.13) peut être réécrit :

$$\begin{cases} \alpha^* = \max_{\alpha, \lambda, \nu} \alpha \\ \text{s.c.} \quad \nu = f(x_j) + g_j^T(-\alpha \sum_{i \in \tilde{I}} \lambda_i g_i + y - x_j), j \in \tilde{I} \\ \nu \geq f(x_j) + g_j^T(-\alpha \sum_{i \in \tilde{I}} \lambda_i g_i + y - x_j), j \notin \tilde{I} \\ \sum_{i \in \tilde{I}} \lambda_i = 1 \\ \alpha \geq 0 \quad \lambda_i \geq 0, i \in \tilde{I}. \end{cases} \quad (3.14)$$

Le problème (3.14) n'a pas une structure linéaire et a une région admissible non vide. Une fois que sa solution $(\alpha^*, \lambda^*, \nu^*)$ a été trouvée, si $\alpha^* > \tilde{\alpha}$ le point

$x(\alpha^*)$ peut s'exprimer

$$\begin{aligned}
x(\alpha^*) &= -\alpha^* \sum_{i \in \tilde{I}} \lambda_i^* g_i + y \\
&= -\alpha^* \sum_{i \in \tilde{I}} \lambda_i^* g_i + \tilde{x} + \tilde{\alpha} \hat{g}(\tilde{\alpha}) \\
&= \tilde{x} - (\alpha^* - \tilde{\alpha}) \left[\frac{\alpha^*}{\alpha^* - \tilde{\alpha}} \sum_{i \in \tilde{I}} \lambda_i^* g_i - \frac{\tilde{\alpha}}{\alpha^* - \tilde{\alpha}} \hat{g}(\tilde{\alpha}) \right]
\end{aligned}$$

cette relation signifie que l'on passe de \tilde{x} vers $x(\alpha^*)$ le long d'une direction qui est l'opposée d'une combinaison affine appropriée de sous-gradients de \hat{f} en \tilde{x} .

A la solution optimale du problème (3.14) (pas 2 de l'algorithme), un nouveau "coin" de la trajectoire proximale est atteint. On peut avoir, soit que certaines contraintes des contraintes d'inégalité

$$\nu \geq f(x_j) + g_j^T(-\alpha \sum_{i \in \tilde{I}} \lambda_i g_i + y - x_j), \quad j \notin \tilde{I} \quad (3.15)$$

sont actives, soit qu'elles satisfont toutes une inégalité stricte. En fait, on a $I_{x(\alpha^*)} \supseteq \tilde{I}$, avec une inclusion stricte dans le premier cas (puisque l'on a plus d'indices actifs dans $I_{x(\alpha^*)}$ que dans \tilde{I} dans ce cas) et une égalité dans le second cas (puisque les contraintes ne sont pas actives, $I_{x(\alpha^*)}$ ne contient pas plus d'éléments que \tilde{I}). En particulier, si on a l'inclusion stricte, on retourne résoudre le problème (3.14) au pas 2, avec un plus grand ensemble actif. Sinon, on a l'égalité et on va au pas 5, où on doit encore résoudre un problème du type (3.14), mais seulement après avoir détecté l'ensemble actif relatif aux points $x(\alpha)$, $\alpha > \alpha^*$, proches de $x(\alpha^*)$. Un tel ensemble actif, disons $I_{x(\alpha)}$, doit être un sous-ensemble propre de $I_{x(\alpha^*)}$. Donc, pour implémenter le pas 5 de l'algorithme 3.4, il est nécessaire de résoudre de façon répétitive le problème (3.14) pour différentes configurations de l'ensemble actif, jusqu'à ce que l'on ait une solution avec une valeur optimale de α^* plus grande que celle actuelle.

Le cas $\alpha^* = 0$ peut arriver seulement si $\tilde{x} = y$. Si on a cela au pas 2, l'algorithme saute directement au pas 5 où on doit encore détecter un sous-ensemble propre approprié de l'ensemble actif.

La proposition suivante est utile afin de sélectionner le sous-ensemble de l'ensemble actif.

Proposition 3.3.1 *Si à la solution optimale de (3.14), aucune contrainte d'inégalité (3.15) n'est active, et les sous-gradients $\{g_j\}$, $j \in \tilde{I}$, sont linéairement indépendants, alors les sous-gradients g_j tels que $\lambda_j^* = 0$ n'appartiennent pas à $I_{x(\alpha^*)}$, pour $\alpha > \alpha^*$, avec $x(\alpha)$ proche de $x(\alpha^*)$.*

Preuve : Puisqu'aucune contrainte d'inégalité n'est active, l'ensemble $I_{x(\alpha)}$ est un sous-ensemble propre de $I_{x(\alpha^*)}$, pour $\alpha > \alpha^*$, avec $x(\alpha)$ proche de $x(\alpha^*)$.

Le point $x(\alpha)$, $\alpha > \alpha^*$ satisfait la condition

$$x(\alpha) - y = -\alpha \hat{g}(x(\alpha)),$$

où

$$\hat{g}(x(\alpha)) \triangleq \sum_{i \in I_{x(\alpha)}} \lambda_i g_i \in \partial \hat{f}(x(\alpha)),$$

pour un choix approprié des multiplicateurs $\lambda_i \geq 0$ tels que $\sum_{i \in I_{x(\alpha)}} \lambda_i(\alpha) = 1$.

Par la continuité de la trajectoire $x(\alpha)$ et comme le sous-différentiel est fermé (car compact), on a qu'il existe un ensemble de multiplicateurs $\lambda'_i \geq 0$, $i \in I_{x(\alpha)}$, avec $\sum_{i \in I_{x(\alpha)}} \lambda'_i = 1$ tels que

$$\hat{g}(x(\alpha^*)) = \lim_{\alpha \searrow \alpha^*} \hat{g}(x(\alpha)) = \sum_{i \in I_{x(\alpha)}} \lambda'_i g_i.$$

D'un autre côté, on a que

$$\hat{g}(x(\alpha^*)) = \sum_{i \in I_{x(\alpha^*)}} \lambda_i^* g_i,$$

avec $\sum_{i \in I_{x(\alpha^*)}} \lambda_i^* = 1$, $\lambda_i^* \geq 0$, $i \in I_{x(\alpha^*)}$.

Et par l'hypothèse d'indépendance linéaire, on doit avoir que $\lambda_i^* = 0$, $\forall i \in I_{x(\alpha^*)} \setminus I_{x(\alpha)}$.

En effet, on a que

$$\begin{aligned}
& \sum_{i \in I_x(\alpha)} \lambda'_i g_i = \sum_{i \in I_x(\alpha^*)} \lambda_i^* g_i \\
\iff & \sum_{i \in I_x(\alpha^*)} \lambda_i^* g_i - \sum_{i \in I_x(\alpha)} \lambda'_i g_i = 0 \\
\iff & \sum_{i \in I_x(\alpha)} (\lambda_i^* - \lambda'_i) g_i + \sum_{i \in I_x(\alpha^*) \setminus I_x(\alpha)} \lambda_i^* g_i = 0 \\
\implies & \lambda_i^* = 0, \quad i \in I_x(\alpha^*) \setminus I_x(\alpha). \quad \square
\end{aligned}$$

Comme il a été remarqué auparavant, le programme (3.14) n'est pas linéaire. Il peut être transformé en programme linéaire équivalent chaque fois que des valeurs positives de α sont admissibles. Comme on l'a mentionné, cela ne pouvait pas se passer uniquement au passage au pas 2, quand $\tilde{x} = y$ car alors $\alpha = 0$. Cependant, près de $\tilde{x} = y$, la trajectoire proximale coïncide avec la direction de plus forte pente de \hat{f} en x . On suppose donc que des valeurs positives de α sont admissibles dans le problème (3.14).

En divisant les deux premiers groupes de contraintes par α , en posant $\beta \triangleq 1/\alpha$ et $\delta \triangleq \nu/\alpha = \nu\beta$, on obtient le programme linéaire suivant :

$$\left\{ \begin{array}{l} \min_{\beta, \lambda, \delta} \quad \beta \\ \text{s.c.} \quad g_j^T \sum \lambda_i g_i - \beta f_j(y) + \delta = 0, \quad j \in \tilde{I} \\ \quad \quad g_j^T \sum_{i \in \tilde{I}} \lambda_i g_i - \beta f_j(y) + \delta \geq 0, \quad j \notin \tilde{I} \\ \quad \quad \sum_{i \in \tilde{I}} \lambda_i = 1 \\ \quad \quad \beta \geq 0 \quad \lambda_i \geq 0, \quad i \in \tilde{I}, \end{array} \right. \quad (3.16)$$

équivalent au problème (3.14) et entièrement défini par l'ensemble \tilde{I} . En indiquant par β^* , δ^* et λ_i^* , $i \in \tilde{I}$, la solution optimale du problème (3.16), il est facile de vérifier que $\beta^* = 0$ correspond au cas où la trajectoire proximale émanant de \tilde{x} se réduit à la demi-droite définie par $\{x \mid x = \tilde{x} = \gamma d^*, \gamma > 0\}$, où $d^* \triangleq - \sum_{i \in \tilde{I}} \lambda_i^* g_i$, ou au cas où le point courant est le point optimal de \hat{f} ,

chaque fois que $d^* = 0$ (car alors $0 \in \sum_{i \in \tilde{I}} \lambda_i^* g_i \in \partial \hat{f}(\tilde{x})$ et donc la condition d'optimalité est vérifiée).

Le problème (3.16) n'est pas admissible seulement quand $\tilde{x} = y$ (ce cas correspond au cas où $\alpha^* = 0$ dans le problème (3.14), avec un saut au pas 5

comme conséquence). Une fois que (3.16) a été résolu, on peut obtenir immédiatement la solution optimale de (3.14).

Le problème (3.16) a une structure *LP*. Concentrons-nous maintenant sur sa solution. Il peut être écrit sous la forme compactée suivante :

$$\begin{cases} \min_{\beta, \lambda, \delta} & \beta \\ \text{s.c.} & \tilde{G}^T \tilde{G} \lambda - \tilde{b} \beta + \tilde{e} \delta = 0 \\ & \tilde{e}^T \lambda = 1 \\ & \hat{G}^T \tilde{G} \lambda - \hat{b} \beta + \hat{e} \delta \geq 0 \\ & \beta \geq 0 \quad \lambda \geq 0, \end{cases} \quad (3.17)$$

où

- \tilde{G} et \hat{G} sont les matrices dont les colonnes sont respectivement les vecteurs g_j , $j \in \tilde{I}$, et g_j , $j \notin \tilde{I}$,
- \tilde{b} et \hat{b} sont les vecteurs dont les composantes sont respectivement les $f_j(y)$, $j \in \tilde{I}$, et les $f_j(y)$, $j \notin \tilde{I}$,
- \tilde{e} et \hat{e} sont des vecteurs de dimensions appropriées,
- λ est le vecteur des λ_i de dimension $|\tilde{I}|$.

On observe que en général, tenant compte des contraintes d'égalité, les variables λ et δ peuvent être exprimées comme fonction de la variable scalaire β , pourvu que la matrice

$$C \triangleq \begin{pmatrix} \tilde{G}^T \tilde{G} & \tilde{e} \\ \tilde{e}^T & 0 \end{pmatrix}$$

soit non singulière. Le problème (3.16) se réduit donc à trouver le plus petit scalaire non négatif β^* , qui satisfait les inégalités du problème (3.17). Notons que la matrice C est non singulière chaque fois que les colonnes de \tilde{G} sont linéairement indépendantes.

D'un autre côté, comme le problème (3.16) est admissible pour $\beta = \tilde{\beta} > 0$, il est facile de vérifier que chaque fois qu'une colonne de \tilde{G} , disons g^r , dépend linéairement des colonnes restantes, i.e.

$$g^r = \sum_{j \neq r} \mu_j g_j, \quad \text{avec} \quad \sum_{j \neq r} \mu_j = 1,$$

on doit avoir la relation suivante :

$$f_r(y) = \sum_{j \neq r} \mu_j f_j(y). \quad (3.18)$$

En effet, comme le problème (3.16) est admissible pour $\tilde{\beta}$, on a que $\tilde{\beta}$ vérifie la contrainte

$$\begin{aligned} & g_r^T \left(\sum_{i \in \tilde{I}} \lambda_i g_i \right) - \tilde{\beta} f_r(y) + \delta = 0 \\ \iff & \left(\sum_{j \neq r} \mu_j g_j \right)^T \left(\sum_{i \in \tilde{I}} \lambda_i g_i \right) - \tilde{\beta} f_r(y) + \delta = 0. \end{aligned}$$

Mais on a aussi, en sommant sur tous les indices $j \neq r$, $j \in \tilde{I}$, puis en multipliant par μ_j :

$$\left(\sum_{j \neq r} \mu_j g_j \right)^T \left(\sum_{i \in \tilde{I}} \lambda_i g_i \right) - \tilde{\beta} \sum_{j \neq r} \mu_j f_j(y) + \delta = 0.$$

On en conclut donc que

$$f_r(y) = \sum_{j \neq r} \mu_j f_j(y) + \delta.$$

La conséquence de la relation (3.18) est que la colonne g_r peut être retirée de \tilde{G} .

Donc, après suppression d'une colonne ou plus de \tilde{G} , on a toujours une matrice C non singulière.

L'inversion de la matrice C mérite quelques observations. Dans certains cas (par exemple chaque fois que la dimension du faisceau est dominante par rapport au nombres de variables), une technique appropriée est de calculer la factorisation QR de \tilde{G} . Si \tilde{G} est de rang plein, alors il existe une formule bien connue pour exprimer C^{-1} .

Chaque fois que pendant le calcul de la factorisation, \tilde{G} est de rang déficient, il est possible de passer à la factorisation QR de la matrice

$$\begin{pmatrix} \tilde{G} \\ e^T \end{pmatrix}, \quad (3.19)$$

en utilisant une technique de mise à jour du type décrit dans [13]. Puisque les colonnes de \tilde{G} sont linéairement indépendantes ssi la matrice (3.19) est de rang plein, on peut conclure de sa factorisation QR que, soit il existe au moins une colonne g_r à supprimer (cas où les colonnes sont linéairement dépendantes), soit on obtient sa factorisation QR qui peut être utilisée dans le calcul de l'inverse de C .

3.4 La méthode

Retournons maintenant au problème (3.1). Dans les méthodes faisceaux classiques, le centre de stabilité est mis à jour au point courant quand ce dernier satisfait un certain test de descente suffisante sur la fonction objectif. De plus, à chaque itération on résout un programme quadratique du type de (3.3).

- Les caractéristiques principales de l'approche que nous allons voir sont :
- mettre chaque fois à jour le centre de stabilité au point généré par l'algorithme quand la fonction polyédrale fournit une approximation suffisamment précise de la fonction f (on verra par la suite que cela correspond à une condition de descente suffisante également) ;
 - sélectionner automatiquement le paramètre de poids α , pourvu qu'il appartienne à un intervalle qui convient $[\alpha_{min}, \alpha_{max}]$.

Etant donné un faisceau B et un centre de stabilité y , on définit un "coin" de la trajectoire proximale par tous les points $x_s \triangleq x(\alpha_s)$, où $\alpha_s > 0$ est un maximiseur d'un programme du type (3.14), pour $s = 1, 2, \dots$. Les coins sont calculés au moyen de la méthode décrite dans la section précédente (algorithme 3.4).

Avant de donner une description formelle de notre méthode, on signale que l'itération générique de l'algorithme est composée de deux phases. L'information principale est le centre de stabilité courant et le faisceau. Dans la première phase, la trajectoire proximale de la fonction polyédrale émanant du centre de stabilité est construite jusqu'à ce que l'on rencontre une condition convenable sur la valeur de α . Ici, à part le besoin d'une sauvegarde appropriée sur les valeurs de α , il n'y a pas besoin d'un choix de départ pour α .

A la sortie de la première phase, la valeur de α , disons $\bar{\alpha}$, est connue et, pour une telle valeur fixée, le problème (3.3) est résolu dans la seconde phase pour des modèles polyédraux de plus en plus précis, jusqu'à ce qu'une condition établie conventionnellement soit satisfaite entre les valeurs de la fonction objectif et celles de l'approximation polyédrale.

A ce moment-ci, à moins d'avoir rencontré un critère d'arrêt, le centre de stabilité est mis à jour et on fait une nouvelle itération principale.

L'algorithme exige les paramètres d'entrée suivants :

- un paramètre de test d'arrêt $\delta > 0$;
- un paramètre de précision m , $0 < m < 1$;
- un paramètre de réduction m_1 , $0 < m_1 < 1$;
- un point de départ $x_1 \in \mathbb{R}^n$, un centre de stabilité initial $y = x_1$ et un faisceau initial $B = \{(y, f(y), g)\}$, où $g \in \partial f(y)$;
- une sauvegarde $0 < \alpha_{max} < \frac{1}{m\delta}$ sur les valeurs de α .

Décrivons maintenant l'itération principale générique, i.e. les pas entre deux mises à jour successives du centre de stabilité. Pour des raisons de lisibilité, on n'introduit pas d'indice d'itération principale.

Itération principale

Phase 1

1. Mettre $\alpha_{min} := \min\{\alpha_{max}, \frac{m_1(1+|f(y)|)}{\|g\|^2}\}$;
2. pour $s = 1, 2, \dots$ et pour des valeurs croissantes de α_s , trouver le premier coin $x_{\bar{s}} = x(\alpha_{\bar{s}})$ tel que

$$\alpha_{min} \geq \alpha_{\bar{s}} \geq \alpha_{max} \quad (3.20)$$

et

$$f(x_{\bar{s}}) - \hat{f}(x_{\bar{s}}) > m\alpha_{\bar{s}}\|\hat{g}(\alpha_{\bar{s}})\|^2, \quad (3.21)$$

où $\hat{g}(\alpha_{\bar{s}})$ est le sous-gradient de \hat{f} au point $x_{\bar{s}}$ satisfaisant la condition d'optimalité (3.8).

3. Si aucun coin $x_{\bar{s}}$ satisfaisant (3.20) et (3.21) n'existe, alors mettre $\bar{\alpha} := \alpha_{max}$ et calculer $x(\bar{\alpha})$ et $f(x(\bar{\alpha}))$, sinon mettre $\bar{\alpha} := \alpha_{\bar{s}}$. Sortir de la phase 1 et entrer dans la phase 2 de l'itération.

Phase2

1. Mettre $k := 1$ et $x_k := x(\bar{\alpha})$. Calculer $g_k \in \partial f(x_k)$. Mettre $B := B \cup \{(x_k, f(x_k), g_k)\}$.
2. Pour la fonction polyédrale \hat{f} , calculer le point $x_{k+1} = x(\bar{\alpha})$ et le sous-gradient \hat{g}_{k+1} qui lui correspond. Calculer également $f(x_{k+1})$ et $g_{k+1} \in \partial f(x_{k+1})$.
3. Si

$$f(x_{k+1}) - \hat{f}(x_{k+1}) \leq m\bar{\alpha}\|\hat{g}_{k+1}\|^2 \quad (3.22)$$

aller au pas 4. Sinon, mettre B à jour en y enlevant si possible un triplet ou plus et en y insérant les deux triplets $(x_{k+1}, \hat{f}(x_{k+1}), \hat{g}_{k+1})$ et $((x_{k+1}, f(x_{k+1}), g_{k+1}))$. Mettre $k := k + 1$ et retourner au pas 2.

4. Si $\|\hat{g}_{k+1}\| \leq \delta$ alors STOP : $x^* = x_{k+1}$; sinon sélectionner x_{k+1} comme nouveau centre de stabilité et sortir de l'itération principale.

Remarque 3.4.1 La valeur de α_{\min} , qui bien sûr n'est pas plus grande que α_{\max} , est choisie (pas 1 de la phase 1) afin de permettre, si la fonction \hat{f} coïncide avec la linéarisation prise au centre de stabilité courant y , une variation plus grande en module qu'une fraction m_1 de $|f(y)|$. En effet, par l'équation (3.21), on a :

$$\begin{aligned} f(x_{\bar{s}}) - \hat{f}(x_{\bar{s}}) &> m \frac{m_1(1+|f(y)|)}{\|g\|^2} \underbrace{\|\hat{g}(\alpha_{\bar{s}})\|^2}_{=\|g\|^2 \text{ car } \hat{f} \text{ coïncide avec la linéarisation prise en } y} \\ \iff f(x_{\bar{s}}) - \hat{f}(x_{\bar{s}}) &> \underbrace{mm_1}_{>0} + mm_1|f(y)| \\ \iff f(x_{\bar{s}}) - \hat{f}(x_{\bar{s}}) &> mm_1|f(y)| \end{aligned}$$

Le lemme suivant est utile pour montrer que la condition (3.22) correspond à une condition de descente suffisante du type

$$f(x_{k+1}) - f(y) \leq (m-1)\bar{\alpha}\|\hat{g}_{k+1}\|^2. \quad (3.23)$$

Lemme 3.4.1 Soient $x_{k+1} \triangleq x(\bar{\alpha}, \hat{f}, y)$ et $g_{k+1} \in \partial \hat{f}(x_{k+1})$ tels que

$$x_{k+1} = y - \bar{\alpha}\hat{g}_{k+1}. \quad (3.24)$$

Si pour un $\epsilon > 0$ donné

$$f(x_{k+1}) - \hat{f}(x_{k+1}) \leq \epsilon, \quad (3.25)$$

alors

$$\hat{g}_{k+1} \in \partial_{\epsilon} f(x_{k+1})$$

et

$$f(x_{k+1}) - f(y) \leq -\bar{\alpha}\|\hat{g}_{k+1}\|^2 + \epsilon. \quad (3.26)$$

Preuve : Par définition de \hat{g}_{k+1} ,

$$\hat{f}(x) \geq \hat{f}(x_{k+1}) + \hat{g}_{k+1}^T(x - x_{k+1}) \quad \forall x \in \mathbb{R}^n.$$

Puisque $f(x) \geq \hat{f}(x) \quad \forall x \in \mathbb{R}^n$ (car \hat{f} est une approximation par le bas de f), on a en tenant compte de (3.25) :

$$f(x) \geq f(x_{k+1}) + \hat{g}_{k+1}^T(x - x_{k+1}) - \epsilon, \quad \forall x \in \mathbb{R}^n. \quad (3.27)$$

L'inégalité (3.26) est obtenue facilement en mettant $x = y$ dans (3.27) et en tenant compte de (3.24). \square

La condition (3.23) est facilement obtenue de (3.26) en sachant que dans l'algorithme, le rôle de ϵ est joué par la quantité $m\bar{\alpha}\|\hat{g}_{k+1}\|^2$. En outre, il est facile de vérifier que la satisfaction de la condition (3.23) implique que

$$f(x_{k+1}) - f(y) \leq (1 - m)\hat{f}'(x_{k+1}; d),$$

où $d = x_{k+1} - y$ et $\hat{f}'(x_{k+1}; d)$ est la dérivée directionnelle de \hat{f} le long de d . En effet,

$$\begin{aligned} f(x_{k+1}) - f(y) &\leq -\hat{g}_{k+1}^T(y - x_{k+1}) + \epsilon \leq \hat{g}_{k+1}^T d + \epsilon \\ &= m\bar{\alpha}\|\hat{g}_{k+1}\|^2 - \bar{\alpha}\|\hat{g}_{k+1}\|^2 \\ &\leq \hat{f}'(x_{k+1}; d) + \epsilon \end{aligned}$$

et comme $\hat{g}_{k+1}^T d = -\bar{\alpha}\|\hat{g}\|^2$, on a que $\epsilon = m\bar{\alpha}\|\hat{g}_{k+1}\|^2 = -m\hat{f}'(x_{k+1}; d)$. Et donc finalement, la relation recherchée.

Par la suite on montrera que, selon une valeur fixée $\bar{\alpha} > 0$ et un centre de stabilité y donné, la condition (3.22) est vérifiée quand le modèle \hat{f} devient de plus en plus précis à mesure que l'on enrichit le faisceau.

En particulier, le 3.4.2 montre que, chaque fois qu'il n'existe pas une bonne conformité entre la fonction et le modèle, la distance entre le point nouvellement généré et ceux restants est borné loin de zéro. Rappelons que, d'après les définitions de la section 2, J est l'ensemble d'indices du faisceau. B .

Lemme 3.4.2 Soit $x_{k+1} \triangleq x(\bar{\alpha}, \hat{f}, y)$. Si

$$f(x_{k+1}) - \hat{f}(x_{k+1}) > \epsilon \quad \text{avec } \epsilon > 0, \quad (3.28)$$

alors

$$\|x_{k+1} - x_j\| > \frac{\epsilon}{2L}, \quad j \in J, \quad (3.29)$$

où $L > 0$ est la constante de Lipschitz de f sur une sphère contenant x_{k+1} et les points x_j , $j \in J$.

Preuve : Par définition de \hat{f} , $\forall x \in \mathbb{R}^n$ on a :

$$\hat{f}(x) \geq f_j(x) = f(x_j) + g_j^T(x - x_j), \quad j \in J,$$

où $g_j \in \partial f(x_j)$. Alors, tenant compte de (3.28) et comme $f(x) \geq \hat{f}(x)$, $\forall j \in J$ on a :

$$\begin{aligned} f(x_{k+1}) - \epsilon > \hat{f}(x_{k+1}) &\geq f(x_j) + g_j^T(x_{k+1} - x_j) \\ &\geq \hat{f}(x_j) + \hat{g}_j^T(x_{k+1} - x_j) \\ &\geq f(x_{k+1}) + g_{k+1}^T(x_j - x_{k+1}) + g_j^T(x_{k+1} - x_j), \end{aligned}$$

i.e.

$$\epsilon < (g_{k+1} - g_j)^T(x_{k+1} - x_j).$$

d'où

$$\epsilon < \|g_{k+1} - g_j\| \|x_{k+1} - x_j\| \leq (\|g_{k+1}\| + \|g_j\|) \|x_{k+1} - x_j\|. \quad (3.30)$$

De plus, $\|g_j\| \leq L \quad \forall j \in J \cup \{k+1\}$. En effet, prenons x_j , $j \in J \cup \{k+1\}$ et $g_j \in \partial f(x_j)$. Par définition du sous-différentiel, on a :

$$\begin{aligned} f(x_j + \frac{g_j}{\|g_j\|}) &\geq f(x_j) + g_j^T((x_j + \frac{g_j}{\|g_j\|}) - x_j) = f(x_j) + \|g_j\| \\ \iff \|g_j\| &\leq f(x_j + \frac{g_j}{\|g_j\|}) - f(x_j) \leq L. \end{aligned}$$

Par (3.30) et comme $\|g_j\| \leq L \quad \forall j \in J \cup \{k+1\}$, on obtient que

$$\epsilon < 2L \|x_{k+1} - x_j\|. \quad \square$$

Dans notre algorithme, au pas 3 de la phase 2, on adopte la technique d' *agrégation* introduite par Kiwiel [19] pendant la mise à jour du faisceau (voir aussi [22] pour des variantes possibles). Cette technique a été conçue pour prendre en compte le fait que l'implémentation pratique des méthodes faisceaux demande d'assigner un emmagasinage fini dans le faisceau. Les propriétés de convergence doivent être prouvées sous l'hypothèse que le nombre d'éléments dans le faisceau ne peut pas grandir infiniment.

Expliquons maintenant comment fonctionne la procédure dans la phase 2 de l'algorithme. On indicera par k la fonction polyédrale \hat{f} et l'ensemble correspondant J à l'itération k de la phase 2. Soit $x_{k+1} = x(\bar{\alpha})$, le point généré au pas 2 et soit $\hat{g}_{k+1} \in \partial \hat{f}(x_{k+1})$, le sous-gradient correspondant. On définit la *linéarisation agrégée*

$$f^a(x) = \hat{f}(x_{k+1}) + \hat{g}_{k+1}^T(x - x_{k+1}),$$

prise en x_{k+1} et on choisit J' un sous-ensemble non vide de J_k (correspondant à un redéplacement possible d'un triplet ou plus du faisceau). Si on redéfinit la fonction

$$\hat{f}_k(x) \triangleq \max[\max_{j \in J'} \{f(x_j) + g_j^T(x - x_j)\}, f^a(x)],$$

on observe que le point x_{k+1} est encore le minimum de la fonction

$$\phi_{\bar{\alpha}}^k(x) \triangleq \bar{\alpha} \hat{f}_k(x) + \frac{1}{2} \|x - y\|^2.$$

Si en plus, on insert le triplet $(x_{k+1}, f(x_{k+1}), g_{k+1})$ dans le faisceau (comme on le fait dans le pas 3), le nouveau modèle polyédral \hat{f}_{k+1} est défini par

$$\hat{f}_{k+1}(x) = \max\{\hat{f}_k(x), f(x_{k+1}) + g_{k+1}^T(x - x_{k+1})\}.$$

Il est clair que la suite $\{\sigma_k\}$ de valeurs optimales de $\phi_{\bar{\alpha}}^k$ est croissante et bornée supérieurement ($\phi_{\bar{\alpha}}^k \leq \hat{f}(y) \leq f(y)$) et donc convergente par Cauchy-Schwartz.

Lemme 3.4.3 *Pour un $\bar{\alpha} > 0$ donné et un centre de stabilité y , soit $\{x_{k+1}\}$, une suite de points tels que $x_{k+1} \triangleq x(\bar{\alpha}, \hat{f}_k, y)$ calculés au pas 2 de la phase 2. Alors*

$$\rho_k f(x_{k+1}) - \hat{f}_k(x_{k+1}) \searrow 0. \quad (3.31)$$

Preuve : La suite $\{\sigma_k\}$, où

$$\sigma_k \triangleq \bar{\alpha} \hat{f}_k(x_{k+1}) + \frac{1}{2} \|x_{k+1} - y\|^2,$$

est croissante et bornée supérieurement, et donc convergente. En particulier, on a

$$\begin{aligned} \sigma_k &= \bar{\alpha} \hat{f}_k(x_{k+1}) + \frac{1}{2} \|x_{k+1} - y\|^2 \\ &= \bar{\alpha} \hat{f}_k(x_{k+1}) - \bar{\alpha} \hat{f}_{k-1}(x_k) + \bar{\alpha} \hat{f}_{k-1}(x_k) + \frac{1}{2} \|x_{k+1} - x_k + x_k - y\|^2 \\ &= \sigma_{k-1} + \bar{\alpha} (\hat{f}_k(x_{k+1}) - \hat{f}_{k-1}(x_k)) + \frac{1}{2} \|x_{k+1} - x_k\|^2 + (x_{k+1} - x_k)^T (x_k - y). \end{aligned} \quad (3.32)$$

Puisque le point x_k satisfait la condition d'optimalité $x_k = y - \bar{\alpha} \hat{g}_k$, pour $\hat{g}_k \in \partial \hat{f}_{k-1}$, et tenant compte de l'inégalité du sous-gradient qui correspond

à \hat{g}_k , on a de (3.32) que

$$\begin{aligned}
\sigma_k &= \sigma_{k-1} + \bar{\alpha}(\hat{f}_k(x_{k+1}) - \hat{f}_{k-1}(x_k)) + \frac{1}{2}\|x_{k+1} - x_k\|^2 + (x_{k+1} - x_k)^T(x_k - y) \\
&\geq \sigma_{k-1} + \bar{\alpha}\hat{g}_k^T(x_{k+1} - x_k) + \frac{1}{2}\|x_{k+1} - x_k\|^2 + (x_{k+1} - x_k)^T(x_k - y) \\
&= \sigma_{k-1} + (x_{k+1} - x_k)^T \underbrace{(\bar{\alpha}\hat{g}_k + x_k - y)}_{=0 \text{ par la condition d'optimalité}} + \frac{1}{2}\|x_{k+1} - x_k\|^2 \\
&= \sigma_{k-1} + \frac{1}{2}\|x_{k+1} - x_k\|^2.
\end{aligned} \tag{3.33}$$

Donc, la convergence de la suite $\{\sigma_k\}$ implique que la suite $\|x_{k+1} - x_k\|$ converge vers zéro.

Supposons maintenant que la suite ρ_k (> 0) ne converge pas vers zéro. Cela devrait impliquer, d'après le 3.4.2, que $\|x_{k+1} - x_k\|$ est borné loin de zéro, ce qui est une contradiction. \square

Finalement, le théorème suivant montre que l'algorithme converge vers une solution estimée du problème (3.1). Ici, y_i est considéré comme le centre de stabilité à la $i^{\text{ème}}$ itération principale. On notera également par α_i et α_{min}^i , les paramètres $\bar{\alpha}$ et α_{min} à la même itération.

Théorème 3.4.1 *L'algorithme s'arrête en un nombre fini d'itérations à un point x_{k+1} satisfaisant les conditions*

$$\|\hat{g}_{k+1}\| \leq \delta \tag{3.34}$$

et

$$\hat{g}_{k+1} \in \partial_\delta f(x_{k+1}) \tag{3.35}$$

Preuve : Supposons que (3.34) n'est jamais satisfaite, i.e.

$$\forall k \quad \|\hat{g}_{k+1}\| > \delta. \tag{3.36}$$

Alors une suite infinie de centres de stabilité, disons $\{y_i\}$, est générée pour le 3.4.3 (car la condition entraînant le STOP n'est pas vérifiée et donc l'algorithme boucle à l'infini). Par (3.23) et tenant compte de (3.36), on a :

$$f(y_{i+1}) - f(y_i) < (m-1)\bar{\alpha}_i\delta^2 \tag{3.37}$$

De plus, pour chaque i , on a $\bar{\alpha}_i \geq \alpha_{min}^i \geq \frac{m_1}{L}$, où L est la constante de Lipschitz de f sur l'ensemble de niveau $S \triangleq \{y \in \mathbb{R}^n \mid f(y) \leq f(y_1)\}$. Donc

on a de (3.37) que

$$f(y_{i+1}) - f(y_i) < \frac{(m-1)m_1\delta^2}{L^2},$$

qui implique

$$f(y_{i+1}) - f(y_1) < i \frac{(m-1)m_1\delta^2}{L^2},$$

ce qui est une contradiction car le membre de gauche est borné par hypothèse et le membre de droite tend vers $-\infty$ quand $i \rightarrow +\infty$.

Une fois que la condition (3.34) est satisfaite pendant une itération, disons \bar{i} , puisque la condition (3.22) est aussi vérifiée (par construction de l'algorithme), on a :

$$f(x_{k+1}) - \hat{f}_k(x_{k+1}) \leq m\bar{\alpha}_i\delta^2 \leq m\alpha_{max}\delta^2 \leq \delta,$$

en prenant $\alpha_{max} \leq \frac{1}{m\delta}$ et la thèse suit du lemme 4.4.2. \square

Chapitre 4

Minimisation de fonctions non différentiables via des plans sécants et le contrôle de proximité

4.1 Introduction

La plupart des méthodes numériques pour résoudre des problèmes d'optimisation non différentiables ont pour but de minimiser des fonctions convexes de plusieurs variables, et l'analyse convexe est en fait la théorie en arrière-plan [5, 26]. Bien que la théorie du gradient généralisé [1] et la théorie des fonctions codifférentiables [3] fournissent un cadre de travail intéressant pour traiter des fonctions non différentiables, non convexes, elles n'ont apparemment pas encore été totalement exploitées du point de vue numérique.

La plupart des algorithmes existants pour l'optimisation non différentiable tombe dans la classe des algorithmes de type sous-gradient et dilatation de l'espace [27], ou des méthodes du type faisceaux [16, 5, 25], ou des algorithmes du type min-max [8, 6] (la convexité n'est pas nécessaire dans cette dernière classe).

En particulier, la famille des méthodes faisceaux est basée sur la méthode du plan sécant, où la convexité de la fonction objectif est l'hypothèse fondamentale. En fait, l'extension de la méthode du plan sécant au cas non convexe n'est pas évidente. Une observation de base est que, en général, des informations du premier ordre ne fournissent plus une approximation par le bas de la fonction objectif, indépendamment de l'hypothèse de non différentiabilité.

Donc, l'optimisation de l'approximation par le plan sécant ne donne pas nécessairement une bonne estimation de la réduction obtenue de la fonction objectif. De plus, un tel modèle pourrait même manquer l'interpolation de la fonction objectif aux points où sa valeur est connue. D'un autre côté, il est clair qu'un nombre d'idées valides dans le contexte convexe non différentiable sont valables aussi dans le traitement du cas non convexe.

Par exemple, des directions de descente obtenues par l'opposé d'une combinaison convexe de gradients, relatifs à des points situés près l'un de l'autre, paraissent souvent jouir aussi de bonnes propriétés de descente pour les fonctions non convexes, spécialement quand les lignes de contour ont une forme de vallée étroite.

Donc, il paraît raisonnable de dire que la minimisation non convexe non différentiable peut bénéficier de l'expérience de l'optimisation convexe, mais les approches valides en optimisation convexe ne peuvent pas être trivialement étendues.

La plupart des auteurs, qui ont étendu les méthodes faisceaux au cas non convexe, ont considéré des modèles affins par morceaux incluant un possible déplacement vers le bas des parties affines [21, 12, 7]. Cependant, la quantité de déplacement paraît d'une certaine manière arbitraire.

Dans ce chapitre, on présente un algorithme itératif qui est encore basé sur des approximations du premier ordre de la fonction objectif.

La différence principale avec les autres méthodes connues est de faire une distinction entre les parties affines qui présentent une sorte de comportement convexe ou non convexe relatif au point courant dans la procédure itérative. En outre, l'utilisation d'un déplacement vers le bas est restreinte à certains cas particuliers.

4.2 Le modèle

Considérons le problème de minimisation sans contraintes suivant :

$$\begin{cases} \min & f(x) \\ \text{s.c.} & x \in \mathbb{R}^n \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ n'est pas nécessairement différentiable.

On suppose que f est localement Lipschitzienne, i.e. Lipschitzienne sur tout ensemble borné. Puisque f est localement Lipschitzienne, alors f est différentiable presque partout (cfr. théorème de Rademacher). Il est connu que, sous ces hypothèses, on définit en chaque x le sous-différentiel

$$\partial f(x) = \text{conv} \{g \mid g \in \mathbb{R}^n, \nabla f(x_k) \rightarrow g, x_k \rightarrow x, x_k \notin \Omega_f\},$$

où Ω_f est l'ensemble (de mesure nulle) où f n'est pas différentiable. Une extension du sous-différentiel est l' ϵ -sous-différentiel de Goldstein $\partial_\epsilon^G f(x)$ défini par

$$\partial_\epsilon^G f(x) = \text{conv} \{ \overset{?}{\partial_\epsilon^G f}(y) \mid \|y - x\| \leq \epsilon \}.$$

On suppose aussi que l'on sait calculer en chaque x la valeur de la fonction objectif et un sous-gradient $g \in \partial f(x)$, i.e. un élément du sous-différentiel.

On arrive maintenant à la description de l'idée de base de la méthode, mettant l'accent sur les différences par rapport aux méthodes façonnées dans le cas convexe. Notons par x_j l'estimation courante du minimum dans une procédure itérative et par g_j un sous-gradient de f en x_j . Le faisceau d'informations disponibles est l'ensemble d'éléments

$$(x_i, f(x_i), g_i, \alpha_i, a_i) \quad i \in I$$

où $x_i, i \in I$, sont les points considérés pendant la procédure, g_i est un sous-gradient de f en x_i , α_i est l'erreur linéarisée entre la valeur actuelle de la fonction objectif en x_j et le développement linéaire généré en x_i et évalué en x_j , i.e.

$$\alpha_i \triangleq f(x_j) - f(x_i) - g_i^T(x_j - x_i)$$

et

$$a_i \triangleq \|x_j - x_i\|.$$

On rappelle que la méthode classique du plan sécant [4, 17] minimise à chaque itération la fonction sécante $f_j(x)$ définie par

$$f_j(x) = \max_{i \in I} \{f(x_i) + g_i^T(x - x_i)\}.$$

La minimisation de $f_j(x)$ peut être mise sous la forme d'un programme linéaire :

$$\begin{cases} \min_{\eta, x} & \eta \\ \text{s.c.} & \eta \geq f(x_i) + g_i^T(x - x_i), \quad i \in I \end{cases}$$

qui est équivalent à résoudre

$$\begin{cases} \min_{v, d} & v \\ \text{s.c.} & v \geq g_i^T d - \alpha_i, \quad i \in I \end{cases}$$

où d est le déplacement à partir de x_j , i.e. $d \triangleq x - x_j$. Par la suite, nous nous référerons au point x_j comme au centre de stabilité.

Il est important de noter que dans le cas non convexe, α_i peut être négatif, car le développement au premier ordre en tout point ne passe pas forcément au-dessous de l'épigraphe de la fonction (dans le cas convexe, on a que $f(x_j) \geq f(x_i) + g_i^T(x_j - x_i)$).

On va donc partitionner l'ensemble I en deux ensembles I_+ et I_- , définis par :

$$I_+ \triangleq \{i \mid \alpha_i \geq 0\} \quad \text{et} \quad I_- \triangleq \{i \mid \alpha_i < 0\}.$$

Les faisceaux définis par les ensembles I_+ et I_- sont caractérisés par des points qui d'une certaine manière manifeste respectivement un "comportement convexe" et un "comportement concave" relativement à x_j . On observe que I_+ n'est jamais vide car au moins l'élément $(x_j, f(x_j), g_j, 0, 0)$ appartient au faisceau.

L'idée de base de l'approche est de traiter différemment les deux faisceaux pendant la construction d'un modèle affine par morceaux.

On définit les fonctions affines par morceaux suivantes :

$$\Delta^+(d) \triangleq \max_{i \in I_+} \{g_i^T d - \alpha_i\}$$

et

$$\Delta^-(d) \triangleq \min_{i \in I_-} \{g_i^T d - \alpha_i\}.$$

En fait, $\Delta^+(d)$ est conçu pour être une approximation de la fonction différence,

$$h(d) \triangleq f(x_j + d) - f(x_j)$$

qui l'interpole en $d = 0$ (puisque l'indice j appartient à I_+).

D'un autre côté, $\Delta^-(d)$ est une approximation localement "pessimiste" de la fonction différence $h(d)$. Comme en $d = 0$, on a $\Delta^+(0) < \Delta^-(0)$, il semble raisonnable de considérer l'approximation $\Delta^+(d)$ importante aussi longtemps que $\Delta^+(d) \leq \Delta^-(d)$.

En d'autres mots, on introduit une sorte de modèle de région de confiance S défini par :

$$S = \{d \mid \Delta^+(d) \leq \Delta^-(d)\}.$$

De plus, on introduit le contrôle de proximité dans l'approche en définissant la "trajectoire proximale" de $\Delta^+(d)$ comme étant la solution optimale d_γ du programme quadratique convexe suivant, paramétrisé en le paramètre scalaire positif γ , où les contraintes assurent que $d \in S$:

$$QP(\gamma) \begin{cases} z_\gamma = \min_{v,d} & \gamma v + \frac{1}{2} \|d\|^2 \\ \text{s.c.} & v \geq g_i^T d - \alpha_i, \quad i \in I_+ \\ & v \leq g_i^T d - \alpha_i, \quad i \in I_- \end{cases}$$

On observe que $z_\gamma \leq 0$, puisque le couple $(v, d) = (0, 0)$ est admissible. On a par conséquent que la valeur optimale de v ne peut être positive. On peut avoir un aperçu intéressant en considérant le dual du problème $QP(\gamma)$. Celui-ci peut s'écrire sous la forme

$$DP(\gamma) \begin{cases} w_\gamma = \min_{\lambda \geq 0, \mu \geq 0} & \frac{1}{2} \|G_+ \lambda - G_- \mu\|^2 + \alpha_+^T \lambda - \alpha_-^T \mu \\ \text{s.c.} & e^T \lambda - e^T \mu = \gamma, \end{cases}$$

où G_+ et G_- sont des matrices dont les colonnes sont, respectivement, les vecteurs g_i , $i \in I_+$, et g_i , $i \in I_-$. De façon analogue, les termes α_i , $i \in I_+$, et α_i , $i \in I_-$, sont groupés dans les vecteurs α_+ et α_- , respectivement.

En effet, le Lagrangien associé au problème $QP(\gamma)$ est

$$L(v, d, \lambda, \mu) = \gamma v + \frac{1}{2} \|d\|^2 + \sum_{i \in I_+} \lambda_i (-v - \alpha_i + g_i^T d) + \sum_{i \in I_-} \mu_i (v + \alpha_i - g_i^T d).$$

La fonction duale $d(\lambda, \mu) = \min_{v,d} L(v, d, \lambda, \mu)$. Pour trouver le minimum de $L(v, d, \lambda, \mu)$, on doit résoudre le système

$$\begin{cases} \nabla_v L(v, d, \lambda, \mu) = & \gamma - \sum_{i \in I_+} \lambda_i + \sum_{i \in I_-} \mu_i = 0 \\ \nabla_d L(v, d, \lambda, \mu) = & d + \sum_{i \in I_+} \lambda_i g_i - \sum_{i \in I_-} \mu_i g_i = 0 \end{cases}$$

Une solution est $v = 0$ et $d = - \sum_{i \in I_+} \lambda_i g_i + \sum_{i \in I_-} \mu_i g_i$. La fonction duale devient

donc

$$\begin{aligned}
d(\lambda, \mu) &= \frac{1}{2} \left\| \sum_{i \in I_-} \mu_i g_i - \sum_{i \in I_+} \lambda_i g_i \right\|^2 - \sum_{i \in I_+} \lambda_i \alpha_i + \sum_{i \in I_+} (\lambda_i g_i)^T \left[\sum_{i \in I_-} \mu_i g_i - \sum_{i \in I_+} \lambda_i g_i \right] + \\
&\quad \sum_{i \in I_-} \mu_i \alpha_i - \sum_{i \in I_-} (\mu_i g_i)^T \left[\sum_{i \in I_-} \mu_i g_i - \sum_{i \in I_+} \lambda_i g_i \right] \\
&= \frac{1}{2} \|G_- \mu - G_+ \lambda\|^2 - \alpha_+^T \lambda + \alpha_-^T \mu - \underbrace{\left[\|G_+ \lambda\|^2 + \|G_- \mu\|^2 - 2 \sum_{i \in I_+} \sum_{i \in I_-} \lambda_i \mu_i g_i^2 \right]}_{\|G_- \mu - G_+ \lambda\|^2} \\
&= \frac{1}{2} \|G_- \mu - G_+ \lambda\|^2 - \alpha_+^T \lambda + \alpha_-^T \mu. \quad \square
\end{aligned}$$

La solution optimale primale (v_γ, d_γ) est liée à la solution optimale duale $(\lambda_\gamma, \mu_\gamma)$ par les formules suivantes :

$$d_\gamma = -G_+ \lambda_\gamma + G_- \mu_\gamma \quad (4.1)$$

$$v_\gamma = -\frac{1}{\gamma} (\|d_\gamma\|^2 + \alpha_+^T \lambda_\gamma - \alpha_-^T \mu_\gamma). \quad (4.2)$$

La dernière formule est tirée de la condition de complémentarité du problème $QP(\gamma)$, i.e.

$$\begin{cases} \lambda_i (-v + g_i^T d - \alpha_i) = 0, & i \in I_+ \\ \mu_i (v - g_i^T d + \alpha_i) = 0, & i \in I_- \end{cases}$$

En sommant la première égalité sur tous les $i \in I_+$ et la seconde égalité sur tous les $i \in I_-$, puis en additionnant ces deux dernières, on obtient la formule (4.2), sachant que $\gamma = \sum_{i \in I_+} \lambda_i - \sum_{i \in I_-} \mu_i$ et par la formule (4.1).

On remarque que la trajectoire proximale émane du centre de stabilité x_j .

Avant de donner une description formelle de l'algorithme, établissons quelques propriétés simples du problème $QP(\gamma)$.

Lemme 4.2.1 Soient $\gamma_1 > \gamma_2 > 0$. Alors on a les relation suivantes :

- (i) $z_{\gamma_1} \leq z_{\gamma_2}$
- (ii) $v_{\gamma_1} \leq v_{\gamma_2}$
- (iii) $\|d_{\gamma_1}\| \geq \|d_{\gamma_2}\|$

Preuve : (i) Par les définitions de z_γ , v_γ , et d_γ et prenant en compte que $\gamma_1 > \gamma_2 > 0$, on a :

$$z_{\gamma_1} = \gamma_1 v_{\gamma_1} + \frac{1}{2} \|d_{\gamma_1}\|^2 \leq \gamma_1 v_{\gamma_2} + \frac{1}{2} \|d_{\gamma_2}\|^2 \leq \gamma_2 v_{\gamma_2} + \frac{1}{2} \|d_{\gamma_2}\|^2 = z_{\gamma_2}.$$

(ii) Supposons $v_{\gamma_1} > v_{\gamma_2}$. Alors, comme $\gamma_1 > \gamma_2$, on a :

$$0 < (\gamma_1 - \gamma_2)(v_{\gamma_1} - v_{\gamma_2}) = \gamma_1 v_{\gamma_1} + \gamma_2 v_{\gamma_2} - (\gamma_1 v_{\gamma_2} + \gamma_2 v_{\gamma_1}).$$

En ajoutant et soustrayant $\frac{1}{2}\|d_{\gamma_1}\|^2 + \frac{1}{2}\|d_{\gamma_2}\|^2$, au membre de droite, on obtient :

$$\begin{aligned} 0 &< \underbrace{[(\gamma_1 v_{\gamma_1} + \frac{1}{2}\|d_{\gamma_1}\|^2) - (\gamma_1 v_{\gamma_2} + \frac{1}{2}\|d_{\gamma_2}\|^2)]}_{\leq 0 \text{ car } (v_{\gamma_1}, d_{\gamma_1}) \text{ est solution optimale}} \\ &+ \underbrace{[(\gamma_2 v_{\gamma_2} + \frac{1}{2}\|d_{\gamma_2}\|^2) - (\gamma_2 v_{\gamma_1} + \frac{1}{2}\|d_{\gamma_1}\|^2)]}_{\leq 0 \text{ car } (v_{\gamma_2}, d_{\gamma_2}) \text{ est solution optimale}}, \end{aligned}$$

ce qui est une contradiction.

(iii) Supposons que $\|d_{\gamma_1}\| < \|d_{\gamma_2}\|$. Alors (ii) entraîne que

$$\gamma_2 v_{\gamma_1} + \frac{1}{2}\|d_{\gamma_1}\|^2 < \gamma_2 v_{\gamma_2} + \frac{1}{2}\|d_{\gamma_2}\|^2 = z_{\gamma_2},$$

ce qui contredit le fait que $(v_{\gamma_2}, d_{\gamma_2})$ soit solution optimale. \square

Lemme 4.2.2 $\forall \gamma > 0$ on a les propriétés suivantes :

- (i) $\|d_\gamma\| \leq 2\gamma\|g_j\|$
- (ii) $z_\gamma \geq -\frac{1}{2}\gamma^2\|g_j\|^2$
- (iii) $|v_\gamma| \geq \frac{1}{2\gamma}\|d_\gamma\|^2$

Preuve : (i) Comme $z_\gamma \leq 0$, on a :

$$(v_\gamma, d_\gamma) \in \mathcal{D} \triangleq \{(v, d) \mid \gamma + \frac{1}{2}\|d\|^2 \leq 0\}.$$

La propriété suit en notant que la fonction objectif de $QP(\gamma)$ est minorée par

$$\gamma g_j^T d + \frac{1}{2}\|d\|^2 \quad \text{car } \alpha_j = 0. \quad (4.3)$$

(ii) La propriété suit par le fait que $-\frac{1}{2}\gamma^2\|g_j\|^2$ est la valeur minimale de la fonction minorisante (4.3). En effet, $d_{\min} = -\gamma g_j$ et en remplaçant dans (4.3), on obtient $-\frac{1}{2}\gamma^2\|g_j\|^2$.

(iii) C'est une conséquence de $z_\gamma \leq 0$.

4.3 L'algorithme

Dans cette section, on décrit un algorithme basé sur la résolution répétitive du problème $QP(\gamma)$, ou de façon équivalente $DP(\gamma)$. Le coeur de l'algorithme est l'itération principale, i.e. l'ensemble de pas où le centre de stabilité reste inchangé.

Il peut y avoir deux sorties de l'itération principale :

1. tout l'algorithme se termine : lorsque qu'une condition approximative de stationarité est satisfaite.
2. on met à jour le centre de stabilité : lorsqu'une condition de descente suffisante est satisfaite.

L'initialisation de l'algorithme demande un point de départ $x_0 \in \mathbb{R}^n$. Le centre de stabilité initial y est initialisé à x_0 . Le faisceau est juste composé d'un élément $(y, f(y), g(y), 0, 0)$, où $g(y) \in \partial f(y)$, de telle façon que I_- est l'ensemble non vide, tandis que I_+ est un singleton. On doit initialiser les paramètres globaux suivants :

- la tolérance de stationarité $\delta > 0$ et la mesure de proximité $\epsilon > 0$;
- le paramètre de descente $m \in (0, 1)$ et le paramètre de section $\rho \in (m, 1)$;
- le paramètre de réduction $r \in (0, 1)$ et le paramètre de croissance $R > 1$.

Ci-dessous, une courte description de l'algorithme :

1. Initialisation
2. Exécution de l'itération principale
3. Mise à jour du faisceau d'informations en fonction du nouveau centre de stabilité et retour en 2.

Dans ce qui suit, on décrit en détails l'itération principale sans l'indicer par simplicité de notation. Bien sûr, le centre de stabilité y sera le centre de stabilité à l'itération courante.

On initialise les paramètres locaux suivants chaque fois qu'on entre dans l'itération principale (ils sont soumis à de possibles modifications durant l'exécution) :

- la mesure de proximité $\theta > 0$
- les paramètres de sauvegarde sur γ : γ_{min} et γ_{max} , $0 < \gamma_{min} < \gamma_{max}$.

On remarque que, en général, l'itération principale maintient le faisceau d'informations (mis à jour) venant d'itérations précédentes. La mise à jour du faisceau est nécessaire car les quantités α_i et a_i dépendent du centre de stabilité.

Algorithme 4.3.1 (Itération Principale)

1. Si $\|g(y)\| \leq \delta$ alors STOP (stationarité atteinte).

Mettre

$$\gamma_{min} := \frac{r\epsilon}{2\|g(y)\|}, \gamma_{max} := R\gamma_{min}, \theta := r\gamma_{min}\delta.$$

2. Construire la trajectoire proximale d_γ pour les valeurs croissantes de γ et choisir $\hat{\gamma}$ égal à la valeur minimale de $\gamma \in (\gamma_{min}, \gamma_{max})$, tels que :

$$f(y + d_\gamma) > f(y) + mv_\gamma$$

si un tel γ existe. S'il n'existe pas, mettre $\hat{\gamma} := \gamma_{max}$. Aller en 4.

3. On enlève les indices des points dont la distance au centre de stabilité est supérieure à ϵ en mettant

$$I_+ := I_+ \setminus \{i \in I_+ \mid a_i > \epsilon\}$$

et

$$I_- := I_- \setminus \{i \in I_- \mid a_i > \epsilon\}.$$

Calculer

$$g^* = \min_{g \in \text{conv}\{g_i \mid i \in I_+\}} \|g\|.$$

Si $\|g^*\| \leq \delta$ alors STOP (stationarité atteinte).

Sinon mettre $\gamma_{max} := \gamma_{max} - r(\gamma_{max} - \gamma_{min})$ et aller en 2.

4. Mettre $\hat{d} := d_{\hat{\gamma}}$, $\hat{v} := v_{\hat{\gamma}}$, $\hat{x} := y + \hat{d}$. Calculer $\hat{g} \in \partial f(\hat{x})$ et mettre

$$\hat{\alpha} := f(y) - f(\hat{x}) + \hat{g}^T \hat{d}.$$

5. (a) Si $\hat{\alpha} < 0$ et $\|\hat{d}\| > \epsilon$ alors insérer l'élément $(\hat{x}, f(\hat{x}), \hat{g}, \hat{\alpha}, \|\hat{d}\|)$ dans le faisceau pour une valeur appropriée de $i \in I_-$, et mettre $\hat{\gamma} := \hat{\gamma} - r(\hat{\gamma} - \gamma_{min})$.
- (b) Sinon, si $\hat{g}^T \hat{d} \geq \rho \hat{v}$ alors insérer l'élément $(\hat{x}, f(\hat{x}), \hat{g}, \max(0, \hat{\alpha}, \|\hat{d}\|))$ dans le faisceau pour une valeur appropriée de $i \in I_+$.

- (c) Sinon, trouver un scalaire $t \in (0, 1)$ tel que $g(t) \in \partial f(y + t\hat{d})$ satisfait la condition

$$g(t)^T \hat{d} \geq \rho \hat{v},$$

et insérer l'élément $(y + t\hat{d}, f(y + t\hat{d}), g(t), \max(0, \alpha_t), t\|\hat{d}\|)$ dans le faisceau pour une valeur appropriée de $i \in I_+$, où $\alpha_t = f(y) - f(y + t\hat{d}) + tg(t)^T \hat{d}$.

6. Si $\|\hat{d}\| \leq \theta$ alors aller en 3. Si

$$f(\hat{x}) \leq f(y) + m\hat{v}$$

alors mettre le nouveau centre de stabilité $y := \hat{x}$ et SORTIR de l'itération principale.

7. Résoudre $QP(\hat{\gamma})$, ou de façon équivalente $DP(\hat{\gamma})$, pour obtenir la solution optimale primale $(v_{\hat{\gamma}}, d_{\hat{\gamma}})$ et la duale $(\lambda_{\hat{\gamma}}, \mu_{\hat{\gamma}})$, et aller en 4.

On se doit de donner quelques explications :

- Le test de stationarité au pas 1 évite d'exécuter l'itération principale si l'on a déjà assez d'informations disponibles pour établir la stationarité de y .
- La construction de la trajectoire proximale au pas 2 peut être discrétisée en résolvant $QP(\gamma)$ de façon répétitive pour des valeurs croissantes de γ , ou en adoptant des techniques du type de celles décrites dans [9].
- L'analyse du test exécuté au pas 3 nous dit que le fait d'avoir un "petit" déplacement (en norme) d_γ correspondant à une "grande" valeur de γ signifie soit qu'un point stationnaire a été atteint, soit que le modèle est incompatible. On différencie ces deux cas en considérant les mesures de distance a_i (suppression du faisceau au pas 3). On observe que le choix de $\hat{\gamma}$ définit implicitement une contrainte sur la norme de $d_{\hat{\gamma}}$ (voir lemme 4.2.2 (i)). D'un autre côté, $\|d_{\hat{\gamma}}\| \leq \theta$ (au pas 6) n'est jamais conséquence du choix d'un $\hat{\gamma}$ trop petit. En fait, on voit que si $\|g(y)\| > \delta$, on a par le lemme 4.2.2 (i) :

$$\|d_{\gamma_{min}}\| \leq 2\gamma_{min}\|g(y)\| \stackrel{\substack{\theta = r\gamma_{min}\delta \\ \text{(cfr. pas 1)}}}{=} \frac{2\|g(y)\|}{r\delta}\theta > \theta.$$

- On remarque que l'insertion d'un indice de faisceau dans I_+ ou I_- au pas 5 n'est pas simplement basé sur le signe de α_i . En fait, dans le cas $\alpha_i < 0$ et $a_i \leq \epsilon$, l'indice i est mis égal à zéro, i.e. la partie affine correspondante est déplacée vers le bas d'une quantité égale $|\alpha_i|$

(voir aussi [7]). Ceci a pour but de laisser tous les éléments de l' ϵ -sous-différentiel de Goldstein contribuer à la construction de l'approximation polyédrique $\Delta^+(d)$, et de garantir que le modèle interpole la fonction objectif en y . En outre, la réduction de $\hat{\gamma}$, chaque fois qu'un indice de faisceau est inséré dans I_- (cfr. pas 5 (a)), a pour but d'éviter que le même point solution \hat{x} ne soit généré une infinité de fois. Pour expliquer le pas (c) du pas 5, on doit voir que le déplacement vers le bas d'une partie affine, quand $\hat{\alpha} < 0$, ne coupe pas toujours le point solution de $QP(\hat{\gamma})$ généré à l'itération précédente. Une condition suffisante pour une telle coupe est $\hat{g}^T \hat{d} \geq \rho \hat{v}$. Si cette condition n'est pas vérifiée, on a recours à une procédure de type recherche linéaire qui nous permet de trouver un point $y + t\hat{d}$, avec $t \in (0, 1)$, satisfaisant $g(t)^T \hat{d} \geq \rho \hat{v}$, où $g(t) \in \partial f(y + t\hat{d})$.

- Finalement, on observe que chaque fois que le centre de stabilité est mis à jour, les paramètres α_i et a_i doivent être aussi mis à jour pour chaque élément du faisceau. Ceci peut avoir comme résultat de changer l'attribution des indices i correspondants de I_+ vers I_- et vice versa.

4.4 La convergence

Dans cette section, on prouve que l'algorithme se termine en un point satisfaisant une condition approximative de stationarité. En particulier, on prouve que pour $\epsilon > 0$ et $\delta > 0$ donnés, il est possible de mettre les paramètres d'entrée tels que, après un nombre fini d'exécutions de l'itération principale, l'algorithme s'arrête au point y qui satisfait la condition

$$\|g^*\| \leq \delta, \quad \text{avec } g^* \in \partial_\epsilon^G f(y).$$

Dans toute cette section, nous faisons les hypothèses suivantes :

H 1 : f est localement Lipschitzienne ;

H 2 : f est faiblement semi-smooth ;

H 3 : l'ensemble $\mathcal{F}_0 = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$ est compact.

Lemme 4.4.1 Soit $\{(v_{\hat{\gamma}}^{(k)}, d_{\hat{\gamma}}^{(k)})\}_{k \in \mathcal{K}}$, une suite générée pendant une seule itération principale, telle que

$$\|d_{\hat{\gamma}}^{(k)}\| > \theta \quad \text{et} \quad f(y + d_{\hat{\gamma}}^{(k)}) - f(y) > mv_{\hat{\gamma}}^{(k)}$$

afin que l'algorithme boucle du pas 4 au pas 7. Alors, on a que :

- (a) Il existe un indice \bar{k} tel que $\forall k \geq \bar{k}$, $k \in \mathcal{K}$, chaque nouvel indice de faisceau est inséré dans I_+ et $\hat{\gamma}$ reste inchangé avec $\hat{\gamma} < \frac{\epsilon}{2\|g(y)\|}$.
- (b) Si $\rho v_{\hat{\gamma}}^{(k)} > g^{(k)T} d_{\hat{\gamma}}^{(k)}$, alors il existe deux scalaires positifs t_1 et t_2 , $0 \leq t_1 < t_2 < 1$, tels que
$$g(t)^T d_{\hat{\gamma}}^{(k)} \geq \rho v_{\hat{\gamma}}^{(k)} \quad \forall t \in [t_1, t_2].$$
- (c) Chaque fois qu'un nouvel indice de faisceau est inséré dans I_+ , la condition
$$g^{(k)T} d_{\hat{\gamma}}^{(k)} \geq \rho v_{\hat{\gamma}}^{(k)}$$
est vérifiée.

Preuve : (a) On voit que chaque fois qu'un indice de faisceau est inséré dans I_- , $\hat{\gamma}$ est réduit (cfr. pas 5(a)). Donc, \bar{k} est le premier indice tel que $\hat{\gamma}$ est réduit en-dessous du seuil $\frac{\epsilon}{2\|g(y)\|}$. En effet, car dans ce cas, par le lemme 4.2.2 (i) on a que $\|d_{\hat{\gamma}}^{(k)}\| \leq \epsilon \quad \forall k \geq \bar{k}$, ce qui implique (cfr. pas 5(a)) l'insertion dans I_+ et pas de réduction de $\hat{\gamma}$.

(b) Observons d'abord que si $\rho v_{\hat{\gamma}}^{(k)} > g^{(k)T} d_{\hat{\gamma}}^{(k)}$, sachant que par hypothèse la condition de descente suffisante n'est pas vérifiée, i.e.

$$f(y + d_{\hat{\gamma}}^{(k)}) - f(y) > mv_{\hat{\gamma}}^{(k)},$$

on a par le théorème des accroissements finis 1.2.2 qu'il existe un scalaire $\bar{t} \in (0, 1)$ tel que

$$\rho v_{\hat{\gamma}}^{(k)} < mv_{\hat{\gamma}}^{(k)} < f(y + d_{\hat{\gamma}}^{(k)}) - f(y) = g(y + \bar{t}d_{\hat{\gamma}}^{(k)})^T d_{\hat{\gamma}}^{(k)},$$

où $g(y + \bar{t}d_{\hat{\gamma}}^{(k)}) \in \partial f(y + \bar{t}d_{\hat{\gamma}}^{(k)})$. De plus, l'existence de tout un intervalle $[t_1, t_2]$, tel que la condition

$$g(t)^T d_{\hat{\gamma}}^{(k)} \geq \rho v_{\hat{\gamma}}^{(k)} \quad \forall t \in [t_1, t_2]$$

soit vérifiée, vient de l'hypothèse **H 2** (voir aussi [23, 24, 7]).

(c) On voit que la condition $g^{(k)T} d_{\hat{\gamma}}^{(k)} \geq \rho v_{\hat{\gamma}}^{(k)}$ est soit assurée par construction ou par le fait qu'on ait

$$g^{(k)T} d_{\hat{\gamma}}^{(k)} \geq g^{(k)T} d_{\hat{\gamma}}^{(k)} - \hat{\alpha} = f(y + d_{\hat{\gamma}}^{(k)}) - f(y) > mv_{\hat{\gamma}}^{(k)} > \rho v_{\hat{\gamma}}^{(k)}$$

chaque fois que $\hat{\alpha} \geq 0$. \square

On peut maintenant prouver que l'itération principale se termine après un nombre fini d'exécution.

Lemme 4.4.2 *L'itération principale se termine après un nombre fini de pas à condition que*

$$\gamma_{\min} < \frac{\epsilon}{2\|g(y)\|}, \quad \theta < \gamma_{\min}\delta, \quad \gamma_{\min} < \gamma_{\max}.$$

Preuve : Pour prouver que l'itération principale se termine, il est nécessaire de démontrer que soit le STOP au pas 3, soit la SORTIE au pas 6 sont réalisés en un nombre fini de pas.

1. Commençons par montrer que l'algorithme ne passe pas une infinité de fois par le pas 3.

Supposons par l'absurde que ce ne soit pas le cas et indiquons par $k \in \mathcal{K}$ toutes les quantités se référant au $k^{\text{ème}}$ passage. On a

$$\|d_{\hat{\gamma}}^{(k)}\| \leq \theta \quad \text{et} \quad \|g^{*k}\| > \delta.$$

Observons que, comme $\hat{\gamma} \leq \gamma_{\max}$, et que par construction γ_{\max} devient plus petit que le seuil $\frac{\epsilon}{2\|g(y)\|}$ en un nombre fini de pas, on a que $\hat{\gamma} < \frac{\epsilon}{2\|g(y)\|}$. Donc, on a asymptotiquement que $\|d_{\hat{\gamma}}^{(k)}\| \leq \epsilon$, qui à son tour entraîne que les indices des nouveaux éléments de faisceau sont insérés asymptotiquement dans I_+ et ne sont jamais enlevés.

En outre, les règles d'insertion dans le faisceau au pas 5 permettent d'insérer un indice dans I_- seulement si $\|d_{\text{gamma}}\| > \epsilon$, et ceci implique que chaque fois que l'on passe par le pas 3, tous les éléments ayant un indice dans I_- sont enlevés.

A partir des considérations faites ci-dessus, tenant compte de (4.1) et de la contrainte $e^T\lambda - e^T\mu = \hat{\gamma}$ dans le problème dual $DP(\hat{\gamma})$, on voit qu'il existe un indice $\bar{k} \in \mathcal{K}$ tel que $\forall k \geq \bar{k}$, la direction $d_{\hat{\gamma}}^{(k)}$ peut être exprimée sous la forme

$$d_{\hat{\gamma}}^{(k)} = -\hat{\gamma}g^{(k)},$$

avec $g^{(k)} \in \text{conv}\{g_i \mid i \in I_+^{(k)}\}$. Mais par hypothèse, $\|d_{\hat{\gamma}}^{(k)}\| \leq \theta$ et

$\|g^{*k}\| > \delta$, donc on devrait avoir

$$\theta \geq \|d_{\hat{\gamma}}^{(k)}\| = \hat{\gamma}\|g^{(k)}\| > \gamma_{\min}\|g^{*k}\| > \frac{\theta}{\delta}\delta = \theta.$$

ce qui est impossible.

2. Il reste à montrer que l'algorithme ne passe pas une infinité de fois par le pas 6.

Supposons par l'absurde que ce ne soit pas le cas et indiquons par $k \in \mathcal{K}$ toutes les quantités se référant au $k^{\text{ème}}$ passage. On a

$$\|d_{\hat{\gamma}}\| > \theta \quad \text{et} \quad f(\hat{x}) > f(y) + m\hat{v}.$$

On voit que par la première partie (a) du lemme 4.4.1, il existe un indice \bar{k} tel que $\forall k \geq \bar{k}$, l'indice de chaque nouvel élément de faisceau est mis dans I_+ avec $\hat{\gamma}$ restant inchangé. Sous une telle condition, pour $k \geq \bar{k}$, la suite $\{z_{\hat{\gamma}}^{(k)}\}$ est croissante (car on augmente les contraintes et on a donc moins de chance de minimiser), borné et donc convergente. De plus, puisque la suite $\{d_{\hat{\gamma}}^{(k)}\}$ est bornée en norme (puisque $\{z_{\hat{\gamma}}^{(k)}\}$ l'est), elle admet une sous-suite convergente, disons $\{d_{\hat{\gamma}}^{(k)}\}_{k \in \mathcal{K}' \subseteq \mathcal{K}}$.

Les considérations ci-dessus impliquent que la suite $\{v_{\hat{\gamma}}^{(k)}\}_{k \in \mathcal{K}' \subseteq \mathcal{K}}$ est aussi convergente ($z_{\gamma} = \min \gamma v + \frac{1}{2}\|d\|^2$) vers une limite négative ($z_{\gamma} \leq 0$), disons \bar{v} . Maintenant supposons que $\bar{v} < 0$, soient s et t deux indices successifs dans \mathcal{K}' et $\beta^{(s)} = \max\{0, \alpha^{(s)}\}$. On a

$$v_{\hat{\gamma}}^{(t)} \geq g^{(s)T} d_{\hat{\gamma}}^{(t)} - \beta^{(s)}, \quad (4.4)$$

$$f(y + d_{\hat{\gamma}}^{(s)}) - f(y) > m v_{\hat{\gamma}}^{(s)} \quad (4.5)$$

et

$$g^{(s)T} d_{\hat{\gamma}}^{(s)} \geq \rho v_{\hat{\gamma}}^{(s)} \quad (4.6)$$

On observe que

$$g^{(s)T} d_{\hat{\gamma}}^{(s)} - \beta^{(s)} \geq \rho v_{\hat{\gamma}}^{(s)} \quad (4.7)$$

En effet, deux cas sont possibles :

- (a) si $\beta^{(s)} = \alpha^{(s)} \triangleq f(y) - f(y + d_{\hat{\gamma}}^{(s)}) + g^{(s)T} d_{\hat{\gamma}}^{(s)}$, sachant que $\rho > m$, on a :

$$g^{(s)T} d_{\hat{\gamma}}^{(s)} - \beta^{(s)} = f(y + d_{\hat{\gamma}}^{(s)}) - f(y) \stackrel{(4.5)}{>} m \underbrace{v_{\hat{\gamma}}^{(s)}}_{\leq 0 \text{ car } z_{\hat{\gamma}} \leq 0} > \rho v_{\hat{\gamma}}^{(s)}.$$

(b) si $\beta^{(s)} = 0$, on a :

$$g^{(s)T} d_{\hat{\gamma}}^{(s)} - \beta^{(s)} \geq \rho v_{\hat{\gamma}}^{(s)} \quad \text{par 4.6.}$$

En combinant (4.4) et (4.7), on obtient que

$$v_{\hat{\gamma}}^{(t)} - \rho v_{\hat{\gamma}}^{(s)} \geq g^{(s)T} (d_{\hat{\gamma}}^{(t)} - d_{\hat{\gamma}}^{(s)})$$

et en passant à la limite :

$$(1 - \rho)\bar{v} \geq 0,$$

ce qui contredit $\bar{v} < 0$. D'où on conclut que $\bar{v} = 0$, mais ceci contredit à son tour le fait que $\|d_{\hat{\gamma}}^{(k)}\| > \theta \quad \forall k \in \mathcal{K}$. \square

Remarque 4.4.1 Puisque $\gamma_{\min} = \frac{r\epsilon}{2\|g(y)\|}$ et $\theta = r\gamma_{\min}\delta$, il suit que

$$\theta \geq \frac{r^2\epsilon\delta}{2L_0}, \quad (4.8)$$

où L_0 est la constante de Lipschitz de f sur l'ensemble \mathcal{F}_0 . En effet, soit $x \in \mathcal{F}_0$ et $x^* \in \partial f(x)$. Par définition du sous-différentiel, on a :

$$\begin{aligned} f(x + \frac{x^*}{\|x^*\|}) &\geq f(x) + x^{*T} (x + \frac{x^*}{\|x^*\|} - x) = f(x) + \|x^*\| \\ \iff \|x^*\| &\leq L_0. \end{aligned}$$

Nous sommes maintenant prêts à prouver que tout l'algorithme se termine.

Théorème 4.4.1 $\forall \epsilon > 0$ et $\forall \delta > 0$, l'algorithme se termine en un nombre fini d'itérations principales en un point satisfaisant la condition approximative de stationarité

$$\|g^*\| \leq \delta \quad \text{avec } g^* \in \partial_\epsilon^G f(y). \quad (4.9)$$

Preuve : La condition approximative de stationarité (4.9) est exactement la condition d'arrêt testée au pas 3 de l'itération principale. Supposons maintenant qu'elle ne soit pas vérifiée pour un nombre fini d'exécutions de l'itération

principale. On a donc alors que la condition de descente est satisfaite une infinité de fois. Soit $y^{(k)}$ le centre de stabilité au $k^{\text{ème}}$ passage par l'itération principale. Alors

$$\begin{aligned}
& \|d_{\hat{\gamma}}^{(k)}\| > \theta^{(k)} \\
& f(y^{(k+1)}) \leq f(y^{(k)}) + mv_{\hat{\gamma}}^{(k)} \\
\text{et } & f(y^{(k+1)}) \leq f(y^{(k-1)}) + m[v_{\hat{\gamma}}^{(k)} + v_{\hat{\gamma}}^{(k-1)}] \\
& \vdots \\
& f(y^{(k+1)}) \leq f(y^{(0)}) + m \sum_{i=0}^k v_{\hat{\gamma}}^{(i)} \\
& \iff f(y^{(k+1)}) - f(y^{(0)}) \leq m \sum_{i=0}^k v_{\hat{\gamma}}^{(i)}.
\end{aligned}$$

En passant à la limite, tenant compte que par (4.8) $\|d_{\hat{\gamma}}^{(i)}\|$ est borné loin de zéro et donc que par le lemme 4.2.2 (iii), $v_{\hat{\gamma}}^{(i)}$ est borné loin de zéro, on a :

$$\lim_{k \rightarrow \infty} f(y^{(k+1)}) - f(y^{(0)}) \leq -\infty,$$

ce qui est impossible puisque f est borné par le bas par les hypothèses **H 1** et **H 3**. \square

4.5 Implémentation pratique

L'algorithme décrit dans la section 3 ne peut pas être directement implémenté car il peut demander un stockage illimité. En fait, l'algorithme ne renferme pas de mécanisme pour contrôler la croissance de la taille du faisceau. Les propriétés de convergence, décrites dans la section 4, sont déduites aussi sous l'hypothèse que la taille du faisceau peut grandir indéfiniment.

Donc, il est nécessaire de considérer explicitement que le faisceau a une taille finie et de montrer que la convergence est conservée sous une telle hypothèse. Une façon possible d'attaquer le problème est d'introduire une technique d'agrégation du type imaginé par Kiwiel [18] et largement utilisé dans les méthodes faisceaux. En particulier, soit \hat{x} le point généré au pas 4 de l'itération principale, obtenu en résolvant $QP(\hat{\gamma})$ ou $DP(\hat{\gamma})$. Si on définit

les quantités agrégées

$$g_+ \triangleq \frac{G_+ \lambda_{\hat{\gamma}}}{e^T \lambda_{\hat{\gamma}}}, \quad \alpha^+ \triangleq \frac{\alpha_+^T \lambda_{\hat{\gamma}}}{e^T \lambda_{\hat{\gamma}}}$$

et, dans le cas $\mu_{\hat{\gamma}} \neq 0$

$$g_- \triangleq \frac{G_- \mu_{\hat{\gamma}}}{e^T \mu_{\hat{\gamma}}}, \quad \alpha^- \triangleq \frac{\alpha_-^T \mu_{\hat{\gamma}}}{e^T \mu_{\hat{\gamma}}},$$

il est facile de vérifier que le problème agrégé

$$QP^a(\hat{\gamma}) \begin{cases} \min_{v,d} & \hat{\gamma}v + \frac{1}{2}\|d\|^2 \\ \text{s.c.} & v \geq g_+^T d - \alpha^+ \\ & v \geq g_i^T d - \alpha_i \quad i \in \bar{I}_+ \\ & v \leq g_-^T d - \alpha^- \\ & v \leq g_i^T d - \alpha_i \quad i \in \bar{I}_- \end{cases}$$

a la même solution optimale $(v_{\hat{\gamma}}, d_{\hat{\gamma}})$ que $QP(\hat{\gamma})$, où \bar{I}_+ et \bar{I}_- sont des sous-ensembles arbitraires de I_+ et I_- respectivement.

Bien sûr, dans le cas $\mu_{\hat{\gamma}} = 0$, la formulation du problème agrégé ne contient pas la contrainte $v \leq g_-^T d - \alpha^-$ et $(v_{\hat{\gamma}}, d_{\hat{\gamma}})$ est encore optimale.

L'introduction de l'agrégation permet en fait d'effacer une partie du faisceau et par conséquent de garder la taille du faisceau bornée. D'un autre côté, l'insertion d'un nouvel élément de faisceau au pas 5 peut être implémenté sans aucune modification. La convergence de l'algorithme n'est pas affectée par le mécanisme de l'agrégation. L'argument clé étant que la monotonie de la suite $\{z_{\hat{\gamma}}^{(k)}\}$, nécessaire dans la preuve du lemme 4.4.2, est encore garantie.

Chapitre 5

Une technique faisceau en minimisation non convexe non différentiable

5.1 Introduction

Les méthodes utilisées couramment pour trouver les minima sans contraintes de fonctions non convexes et non nécessairement différentiables se révèlent généralement être une adaptation de méthodes à l'origine conçues pour traiter le problème convexe correspondant.

Parmi de telles méthodes, on peut mentionner celles dues à Kiwiel, Makëla et Neittaanmäki, Schramm et Zowe [21, 12, 7].

Ceux-ci, grâce à la terminologie utilisée par Schramm et Zowe, combinent l'idée de faisceau, due à Lemaréchal, et l'approche de la région de confiance. La technique du faisceau est à son tour tirée de l'approximation d'un plan sécant, où une approximation affine par morceaux (le modèle) de la fonction objectif est minimisée afin d'obtenir une solution approximée du problème de minimisation original. La qualité du modèle est améliorée à chaque itération et ceci est l'argument clé qui garantit la convergence.

Le modèle est défini comme le maximum point par point d'un ensemble de fonctions affines. La fonction originale, si convexe, est interpolée par le modèle en un ensemble de points. Cet ensemble est constitué d'au moins un point, normalement le meilleur possible, auquel nous nous référerons comme au centre de stabilité.

Si la fonction à minimiser est non convexe, la propriété d'interpolation du modèle peut être oubliée. Donc, afin de continuer avec un modèle défini comme le maximum point par point de fonctions affines, certaines parties affines sont translatées verticalement, autant de fois que nécessaire, afin de conserver l'interpolation au moins au centre de stabilité.

Une autre approche a aussi été imaginée, où le modèle prend explicitement en compte le comportement non convexe de la fonction objectif.

Ici, on étend une telle approche en définissant un modèle du type affiné par morceaux pouvant être mis sous la forme d'une différence de deux convexes (*DC*), i.e. la somme d'une fonction convexe affine par morceaux et d'une fonction concave affine par morceaux. Donc, notre approche bénéficie de quelques idées venant de la théorie des fonctions quasi différentiables (=somme de fonctions convexes et concaves).

En liaison avec notre modèle, on adoptera tout le mécanisme venant des méthodes faisceaux.

5.2 Le modèle

Considérons le problème de minimisation suivant :

$$\begin{cases} \min & f(x) \\ \text{s.c.} & x \in \mathbb{R}^n \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ n'est pas nécessairement convexe, ni différentiable.

On suppose que f est localement Lipschitzienne et donc différentiable presque partout (cfr. théorème de Rademacher). Sous ces hypothèses, on définit en chaque x le sous-différentiel

$$\partial f(x) = \text{conv} \{g \mid g \in \mathbb{R}^n, \nabla f(x_k) \rightarrow g, x_k \rightarrow x, x_k \notin \Omega_f\},$$

où Ω_f est l'ensemble (de mesure nulle) où f n'est pas différentiable. Une extension du sous-différentiel est l' ϵ -sous-différentiel de Goldstein $\partial_\epsilon^G f(x)$ défini par

$$\partial_\epsilon^G f(x) = \text{conv} \{\partial^G f(y) \mid \|y - x\| \leq \epsilon\}.$$

On suppose aussi que

- on sait calculer en chaque x la valeur de la fonction objectif et un sous-gradient $g \in \partial f(x)$, i.e. un élément du sous-différentiel.
- pour un point x_0 quelconque, l'ensemble $\mathcal{F}_0 = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$ est compact.

Décrivons maintenant le modèle adopté dans notre méthode, en se concentrant sur les différences par rapport aux méthodes conçues dans le cas convexe.

On note par x_j , l'estimation courante du minimum dans une procédure itérative (c'est le centre de stabilité) et par g_j , un sous-gradient de $f(x_j)$. Le faisceau d'informations qu'on peut obtenir est l'ensemble d'éléments

$$(x_i, f(x_i), g_i, \alpha_i, a_i) \quad i \in I,$$

où x_i , $i \in I$, sont les points touchés pendant la procédure; g_i est un sous-gradient de $f(x_i)$; α_i est l'erreur linéarisée entre la valeur actuelle de la fonction objectif en x_j , et le développement linéaire généré en x_i et évalué en x_j , i.e.

$$\alpha_i \triangleq f(x_j) - f(x_i) - g_i^T(x_j - x_i)$$

et

$$a_i \triangleq \|x_j - x_i\|.$$

La méthode classique du plan sécant minimise à chaque itération la fonction du plan sécant $f_j(x)$, définie par

$$f_j(x) = \max_{i \in I} \{f(x_i) + g_i^T(x - x_i)\}.$$

La minimisation de $f_j(x)$ peut être mise sous la forme d'un programme linéaire :

$$\begin{cases} \min_{\theta, x} & \theta \\ \text{s.c.} & \theta \geq f(x_i) + g_i^T(x - x_i) \quad i \in I, \end{cases}$$

qui est équivalent à résoudre

$$\begin{cases} \min_{v, d} & v \\ \text{s.c.} & v \geq g_i^T d - \alpha_i \quad i \in I, \end{cases}$$

où d est le déplacement à partir de x_j , i.e. $d = x - x_j$.

Il est important de noter que dans le cas non convexe, α_i peut être négatif, car le développement au premier ordre en n'importe quel point ne passe

pas forcément au-dessous de l'épigraphe de la fonction.

On va donc partitionner l'ensemble I en deux ensembles I_+ et I_- , définis comme suit :

$$I_+ \triangleq \{i \mid \alpha_i \geq 0\} \quad \text{et} \quad I_- \triangleq \{i \mid \alpha_i \leq 0\}.$$

On remarque que les deux ensembles ont une intersection non vide, car par définition, au moins l'élément du faisceau $(x_j, f(x_j), g_j, 0, 0)$ appartient à I_+ et I_- .

Les faisceaux définis par les ensembles I_+ et I_- sont caractérisés par des points qui d'une certaine manière font preuve respectivement d'un "comportement convexe" et d'un "comportement concave" relativement à x_j .

L'idée de base de l'approche est de traiter différemment les deux faisceaux lors de la construction d'un modèle affin par morceaux.

On définit les fonctions affines par morceaux suivantes :

$$\Delta^+(d) \triangleq \max_{i \in I_+} \{g_i^T d - \alpha_i\}$$

et

$$\Delta^-(d) \triangleq \min_{i \in I_-} \{g_i^T d - \alpha_i\}$$

On voit que $\Delta^+(d)$ est convexe, tandis que $\Delta^-(d)$ est concave. En outre, on a que

$$\Delta^+(d) \geq \Delta^-(d) \quad \forall d \tag{5.1}$$

car $I_+ \cap I_- \neq \emptyset$. En effet, par l'absurde, si

$$\begin{aligned} \exists d \text{ tel que } & \Delta^+(d) < \Delta^-(d) \\ \iff & \max_{i \in I_+} \{g_i^T d - \alpha_i\} < \min_{i \in I_-} \{g_i^T d - \alpha_i\}. \end{aligned}$$

Et en particulier, comme $\{i \mid \alpha_i = 0\} \subseteq I_+ \cap I_-$, on a :

$$\max_{i \in I_+} \{g_i^T d\} < \min_{i \in I_-} \{g_i^T d\},$$

ce qui est une contradiction.

Maintenant, pour un choix quelconque du scalaire $p \in (0, 1)$, on peut définir la fonction

$$\Delta_p(d) \triangleq p\Delta^+(d) + (1-p)\Delta^-(d),$$

qui peut être interprétée comme une approximation de la fonction différence

$$h(d) \triangleq f(x_j + d) - f(x_j)$$

obtenue par une moyenne pondérée des deux approximations $\Delta^+(d)$ et $\Delta^-(d)$ qui interpolent toutes les deux $h(d)$ en $d = 0$. En effet,

$$\Delta^+(0) = \max_{i \in I_+} \{-\alpha_i\} = 0 = \min_{i \in I_-} \{-\alpha_i\}.$$

La fonction $\Delta_p(d)$ est la fonction modèle que l'on va adopter pour trouver un pas de descente pour f . C'est une fonction *DC* (différence de deux convexes) et affine par morceaux.

On introduit le contrôle de proximité dans notre approche, afin de définir implicitement une sorte de région de confiance, en ajoutant un terme de pénalité quadratique à $\Delta_p(d)$, la fonction modèle. On arrive donc à la fonction modèle complète

$$f_{p\gamma}(d) \triangleq \gamma \Delta_p(d) + \frac{1}{2} \|d\|^2,$$

où γ (le paramètre de contrôle de proximité) est un scalaire positif. On met en évidence dans la notation le fait que le modèle dépende des deux paramètres scalaire p et γ .

La fonction $f_{p\gamma}(d)$ est coercive (i.e. $\lim_{\|d\| \rightarrow \infty} f_{p\gamma}(d) = \infty$) et, en général, peut admettre plusieurs minima locaux.

Le lemme suivant fournit une borne sur la norme $d_{p\gamma}$.

Lemme 5.2.1 $\forall \gamma > 0$, on a :

$$\|d_{p\gamma}\| \leq 2\gamma \|g_{I_-}\|,$$

$$\text{où } \|g_{I_-}\| \triangleq \max_{i \in I_-} \{\|g_i\|\}.$$

Preuve : Par la définition de $\Delta_p(d)$, tenant compte de (5.1) et sachant que $\alpha_i \leq 0, \forall i \in I_-$, on a :

$$\Delta_p(d) \geq \Delta^-(d) \geq \min_{i \in I_-} g_i^T d \geq -\|g_{I_-}\| \|d\|.$$

Et comme $f_{p\gamma}(d) \triangleq \gamma\Delta_p(d) + \frac{1}{2}\|d\|^2$, on a que :

$$\begin{aligned}
& f_{p\gamma}(d) \geq -\gamma\|g_{I_-}\|\|d\| + \frac{1}{2}\|d\|^2 \\
\iff & f_{p\gamma}(d) - \frac{1}{2}\|d\|^2 \geq -\gamma\|g_{I_-}\|\|d\| \\
\iff & \frac{1}{2}\|d_{p\gamma}\|^2 \leq \gamma\|g_{I_-}\|\|d_{p\gamma}\| + \underbrace{f_{p\gamma}(d_{p\gamma})}_{\leq 0} \\
\implies & \|d_{p\gamma}\|^2 \leq 2\gamma\|g_{I_-}\|\|d_{p\gamma}\|. \quad \square
\end{aligned}$$

Dans notre méthode, le programme fortement convexe suivant joue un rôle significatif :

$$\min_d \gamma\Delta^+(d) + \frac{1}{2}\|d\|^2. \quad (5.2)$$

Celui-ci, en introduisant la variable scalaire v , peut être réécrit comme un programme quadratique de la forme :

$$QP(\gamma) \begin{cases} \min_{v,d} & \gamma v + \frac{1}{2}\|d\|^2 \\ \text{s.c.} & v \geq g_i^T d - \alpha_i \quad i \in I_+. \end{cases}$$

Le dual de $QP(\gamma)$ peut s'écrire :

$$DP(\gamma) \begin{cases} \min_{\lambda \geq 0} & \frac{1}{2}\|G_+\lambda\|^2 + \alpha_+^T \lambda \\ \text{s.c.} & e^T \lambda = \gamma, \end{cases}$$

où G_+ est la matrice dont les colonnes sont les vecteurs g_i , $i \in I_+$; les scalaires α_i , $i \in I_+$, sont groupés dans le vecteur α_+ ; $\lambda \geq 0$ est le vecteur multiplicateur de dimension appropriée.

En effet, le Lagrangien associé au problème $QP(\gamma)$ est

$$L(v, d, \lambda) = \gamma v + \frac{1}{2}\|d\|^2 + \sum_{i \in I_+} \lambda_i (-v - \alpha_i + g_i^T d).$$

La fonction duale $d(\lambda) = \min_{v,d} L(v, d, \lambda)$. Pour trouver le minimum de $L(v, d, \lambda)$, on doit résoudre le système

$$\begin{cases} \nabla_v L(v, d, \lambda) = & \gamma - \sum_{i \in I_+} \lambda_i = 0 \\ \nabla_d L(v, d, \lambda) = & d + \sum_{i \in I_+} \lambda_i g_i = 0 \end{cases}$$

Une solution est $v = 0$ et $d = -\sum_{i \in I_+} \lambda_i g_i$. La fonction duale devient donc

$$\begin{aligned} d(\lambda) &= \frac{1}{2} \left\| \sum_{i \in I_+} \lambda_i g_i \right\|^2 - \sum_{i \in I_+} \lambda_i \alpha_i + \sum_{i \in I_+} (\lambda_i g_i)^T \left[- \sum_{i \in I_+} \lambda_i g_i \right] \\ &= \frac{1}{2} \left\| \sum_{i \in I_+} \lambda_i g_i \right\|^2 - \sum_{i \in I_+} \lambda_i \alpha_i - \left\| \sum_{i \in I_+} \lambda_i g_i \right\|^2 \\ &= -\frac{1}{2} \|G_+ \lambda\|^2 - \alpha_+^T \lambda. \quad \square \end{aligned}$$

La solution primale optimale $(d_\gamma, \Delta^+(d))$ de (5.2) est liée à la solution optimale duale λ_γ de $DP(\gamma)$ par les formules

$$d_\gamma = -G_+ \lambda_\gamma \quad (5.3)$$

$$\Delta^+(d_\gamma) = -\frac{1}{\gamma} (\|d_\gamma\|^2 + \alpha_+^T \lambda_\gamma). \quad (5.4)$$

La dernière formule est tirée de la condition de complémentarité du problème $QP(\gamma)$, i.e.

$$\lambda_i (-v + g_i^T d - \alpha_i) = 0, \quad i \in I_+.$$

En sommant cette égalité sur tous les $i \in I_+$, sachant que $\sum_{i \in I_+} \lambda_i = \gamma$ et par la formule (5.3), on obtient la formule (5.4).

5.3 L'algorithme

Décrivons maintenant l'algorithme. Il est basé sur la solution itérative du problème

$$\min_{d \in \mathbb{R}^n} f_{p\gamma}(d).$$

Comme la fonction $f_{p\gamma}(d)$ est non convexe, on entend ici par “résoudre”, trouver une solution optimale globale. On verra dans l'appendice une discussion sur la façon de trouver une telle solution. Par la suite, on indiquera le centre de stabilité x_j par y afin de souligner son rôle particulier par rapport aux autres points possibles.

De plus, dans l'algorithme on se réfère à l'itération principale comme à l'ensemble d'étapes au cours desquelles le centre de stabilité reste inchangé.

Il peut y avoir deux sorties de l'itération principale :

1. tout l'algorithme se termine : lorsque qu'une condition approximative de stationarité est satisfaite.

2. on met à jour le centre de stabilité : lorsque qu'une condition de descente suffisante est satisfaite.

L'initialisation de l'algorithme demande un point de départ $x_0 \in \mathbb{R}^n$. Le centre de stabilité initial y_0 est initialisé à x_0 . Initialement, on a seulement un élément de faisceau $(y_0, f(y_0), g(y_0), 0, 0)$, où $g(y_0) \in \partial f(y_0)$. L'indice correspondant est mis dans I_+ et I_- , qui sont par conséquent tous les deux un singleton. On doit initialiser les paramètres globaux suivants :

- la tolérance de stationarité $\delta > 0$ et la mesure de proximité $\epsilon > 0$;
- le paramètre de descente $m \in (0, 1)$;
- le paramètre de réduction $r \in (0, 1)$ et le paramètre de croissance $R > 1$;
- le paramètre d'équilibre $p_0 \in (0, 1)$.

Ci-dessous, une courte description de l'algorithme :

1. Initialisation
2. Exécution de l'itération principale
3. Mise à jour du faisceau d'informations en fonction du nouveau centre de stabilité et retour en 2.

Dans la description qui suit, on n'indice pas l'itération principale par question de simplicité de notation. Donc, le centre de stabilité y sera le centre de stabilité à l'itération courante.

On initialise les paramètres locaux suivants chaque fois qu'on entre dans l'itération principale (ils sont soumis à de possibles modifications durant l'exécution) :

- les paramètres de sauvegarde sur γ : γ_{min} et γ_{max} , $0 < \gamma_{min} < \gamma_{max}$;
- le paramètre de seuil de descente $\sigma > 0$, le paramètre de seuil d'insertion dans le faisceau $\phi > 0$ et le paramètre d'approximation $\eta > 0$;
- le paramètre d'équilibre $p = p_0$.

On impose les conditions suivantes sur les paramètres pendant l'itération principale :

$$\gamma_{min} < \frac{\epsilon}{2\|g_{I_-}\|} \quad (5.5)$$

$$\delta \geq \sqrt{\frac{\sigma + \eta}{\gamma_{min}}} \quad (5.6)$$

En particulier, on met initialement $\gamma_{min} = \frac{r\epsilon}{2\|g_{I_-}\|}$, pour que la condition (5.5) soit satisfaite ($g_{I_-} = g(y_0)$ au début de la première itération), et on met

aussi $\gamma_{max} = R\gamma_{min}$.

Les paramètres restants sont initialisés tels que la condition (5.6) soit satisfaite. En particulier, on initialise σ tel que $\delta = \sqrt{\frac{2\sigma}{\gamma_{min}}}$ (i.e. $\sigma = \frac{\gamma_{min}\delta^2}{2}$), et alors on pose $\eta = \sigma$. Finalement, on pose $\phi = \eta = \sigma$.

On remarque que, en général, l'itération principale maintient le faisceau d'informations (mis à jour) d'itérations précédentes (les quantités α_i et a_i sont en fait dépendantes du centre de stabilité et ce dernier ne change pas pendant l'itération principale). Notez aussi que l'allocation de n'importe quel indice d'élément de faisceau dans I_+ ou I_- , ou dans les deux, dépend aussi du centre de stabilité à l'itération courante.

Algorithme 5.3.1 (Itération Principale)

1. Si $\|g(y)\| \leq \delta$ alors STOP (stationarité atteinte).
2. Mettre $\gamma_{min} := \min\{\gamma_{min}, \frac{r\epsilon}{2\|g_{I_-}\|}\}$, $\sigma := \frac{\gamma_{min}\delta^2}{2}$ et $\phi := \eta := \sigma$.
Choisir $\gamma \in (\gamma_{min}, \gamma_{max})$ et calculer $d_{p\gamma}$, un minimum global de $f_{p\gamma}(d)$, et d_γ , le minimum de $QP(\gamma)$.
3. Calculer $\Delta^+(d_{p\gamma})$.
Si $\Delta^+(d_{p\gamma}) \leq -\sigma$ aller en 4. Si $\Delta^+(d_{p\gamma}) - \Delta^+(d_\gamma) > \eta$ alors mettre $p := p + r(1 - p)$ et retourner en 2. Sinon aller en 6.
4. Initialiser $\hat{x} := y + d_{p\gamma}$ et calculer $f(\hat{x})$. Si

$$f(\hat{x}) \leq f(y) + m\Delta^+(d_{p\gamma}) \quad (5.7)$$

alors mettre le nouveau centre de stabilité $y := \hat{x}$ et SORTIR de l'itération principale.

5. Calculer $\hat{g} \in \partial f(\hat{x})$ et initialiser

$$\hat{\alpha} := f(y) - f(\hat{x}) + \hat{g}^T d_{p\gamma}.$$

- (a) Si $\hat{\alpha} \leq -\phi$ et $\|d_{p\gamma}\| > \epsilon$ alors mettre $\gamma := \gamma - r(\gamma - \gamma_{min})$ et retourner en 2.
- (b) Si $\hat{\alpha} \geq \phi$ alors insérer l'élément $(\hat{x}, f(\hat{x}), \hat{g}, \hat{\alpha}, \|d_{p\gamma}\|)$ dans le faisceau pour une valeur appropriée de $i \in I_+$ et retourner en 2.
- (c) Si $0 \leq \hat{\alpha} < \phi$ insérer l'élément $(\hat{x}, f(\hat{x}), \hat{g}, 0, \|d_{p\gamma}\|)$ deux fois dans le faisceau, pour deux valeurs appropriées des indices, l'un appartenant à I_+ et l'autre appartenant à I_- , et retourner en 2.

- (d) sinon trouver un scalaire $t \in (0, 1)$ tel que $g(t) \in \partial f(y + td)$ satisfait la condition

$$g(t)^T d_{p\gamma} \geq m\Delta^+(d_{p\gamma}), \quad (5.8)$$

insérer l'élément $(y + td_{p\gamma}, f(y + td_{p\gamma}), g(t), 0, t\|d_{p\gamma}\|)$ deux fois dans le faisceau, pour deux valeurs appropriées des indices, l'un appartenant à I_+ et l'autre appartenant à I_- , et retourner en 2.

6. Mettre

$$I_+ := I_+ \setminus \{i \in I_+ \mid a_i > \epsilon\}$$

et

$$I_- := I_- \setminus \{i \in I_- \mid a_i > \epsilon\}.$$

Calculer

$$\|g^*\| = \min_{g \in \text{conv} \{g_i \mid i \in I_+ \cup I_-\}}.$$

Si $\|g^*\| \leq \delta$ alors STOP (stationarité atteinte).

Sinon mettre :

$$\gamma_{\max} := \gamma_{\max} - r(\gamma_{\max} - \gamma_{\min}) \quad (5.9)$$

et aller en 2.

A la sortie de l'itération principale, une fois que le centre de stabilité a été mis à jour, le faisceau est lui aussi mis à jour selon la procédure suivante :

Algorithme 5.3.2 (Mise à jour du Faisceau)

1. Calculer les erreurs de linéarisation α_i et les distances a_i , en fonction du nouveau centre de stabilité, $\forall i \in I_+ \cup I_-$.
2. Si $|\alpha_i| \leq \phi$ alors insérer l'élément $(x_i, f(x_i), g_i, 0, a_i)$ deux fois dans le faisceau, pour deux valeurs appropriées des indices, l'un appartenant à I_+ et l'autre à I_- .
3. Si $|\alpha_i| < -\phi$ alors insérer l'élément $(x_i, f(x_i), g_i, \alpha_i, a_i)$ dans le faisceau pour une valeur appropriée de l'indice $i \in I_-$.
4. Si $|\alpha_i| > \phi$ alors insérer l'élément $(x_i, f(x_i), g_i, \alpha_i, a_i)$ dans le faisceau pour une valeur appropriée de l'indice $i \in I_+$.
5. Effacer du faisceau tous les éléments tels que :

$$i \in I_- \quad \text{et} \quad x_i \notin \mathcal{F}_0. \quad (5.10)$$

Remarque 5.3.1 Les paramètres locaux γ_{\min} , σ , ϕ et η sont sujets à de possibles modifications durant l'exécution de l'itération principale car γ_{\min} dépend de $\|g_{I_-}\|$ qui peut changer lors du pas 5.

Le paramètre γ_{\min} est cependant borné loin de zéro. en fait, les règles d'insertion dans le faisceau au pas 5 de l'itération principale et la condition de suppression dans le faisceau (5.10) assure que $\|g_{I_-}\| \leq L_f$, où L_f est la constante de Lipschitz de f sur l'ensemble de points auxquels la distance à \mathcal{F}_0 n'est pas plus grande que ϵ . En effet, soit $x_0 \in \mathcal{F}_0$ et $g \in \partial f(x)$. Par définition du sous-différentiel, on a :

$$\begin{aligned} f(x + \frac{g}{\|g\|}) &\geq f(x) + g^T(x + \frac{g}{\|g\|} - x) = f(x) + \|g\| \\ \iff \|g\| &\leq f(x + \frac{g}{\|g\|}) - f(x) \leq L_f. \end{aligned}$$

Donc, d'un bout à l'autre de l'algorithme, on a que $\gamma_{\min} \geq \bar{\gamma} \triangleq \frac{r\epsilon}{2L_f}$, et par conséquent $\sigma = \phi = \eta \geq \bar{\rho} \triangleq \frac{\bar{\gamma}\delta^2}{2}$.

La valeur du paramètre ϕ utilisé dans la procédure de mise à jour du faisceau est celle que l'on a pour sortir de l'itération principale.

5.4 La convergence

Dans toute cette section, nous faisons les hypothèses suivantes :

H 1 : f est localement Lipschitzienne ;

H 2 : f est faiblement semi-smooth ;

En particulier, **H 2** est une hypothèse technique qui assure qu'au pas 5, de l'itération principale, le problème de trouver le scalaire t satisfaisant la condition (5.8) est bien posé.

Avant de prouver la convergence, on établit le lemme suivant qui assure que, en initialisant le paramètre p suffisamment près de 1, la valeur prise par la fonction Δ^+ en $d_{p\gamma}$, le minimum de $f_{p\gamma}(d)$, est proche de celle prise par Δ^+ en d_γ , le minimum du problème $QP(\gamma)$.

Lemme 5.4.1 Soit donné $\gamma > 0, \forall \eta > 0$ il existe un seuil positif $p_{\max} < 1$ tel que $\forall p \geq p_{\max}$ on a la relation suivante :

$$\Delta^+(d_{p\gamma}) \leq \Delta^+(d_\gamma) + \eta,$$

où d_γ est la solution optimale du problème (5.2).

Preuve : Par définition de $d_{p\gamma}$, on a l'inégalité suivante :

$$\begin{aligned} & \gamma[p\Delta^+(d_{p\gamma}) + (1-p)\Delta^-(d_{p\gamma})] + \frac{1}{2}\|d_{p\gamma}\|^2 \\ & \leq \gamma[p\Delta^+(d_\gamma) + (1-p)\Delta^-(d_\gamma)] + \frac{1}{2}\|d_\gamma\|^2. \end{aligned} \quad (5.11)$$

En outre, la convexité de $\Delta^+(d)$ implique $\forall p \in (0, 1)$ et $\forall u \in \partial\Delta^+(d_\gamma)$:

$$\begin{aligned} & \gamma p\Delta^+(d_{p\gamma}) + \frac{1}{2}\|d_{p\gamma}\|^2 \\ & \geq \gamma p\Delta^+(d_\gamma) + \frac{1}{2}\|d_\gamma\|^2 + (\gamma pu + d_\gamma)^T(d_{p\gamma} - d_\gamma) + \frac{1}{2}\|d_{p\gamma} - d_\gamma\|^2. \end{aligned} \quad (5.12)$$

Maintenant, soit $w(d) \triangleq \gamma\Delta^+(d) + \frac{1}{2}\|d\|^2$ la fonction objectif du problème $QP(\gamma)$, et soit $w'(d_\gamma, d_{p\gamma} - d_\gamma)$ la dérivée directionnelle de w évaluée en d_γ le long de la direction $(d_{p\gamma} - d_\gamma)$. L'optimalité de d_γ implique que $w'(d_\gamma, d_{p\gamma} - d_\gamma) \geq 0$. De plus, comme

$$\begin{aligned} w'(d_\gamma, d_{p\gamma} - d_\gamma) &= \max_{\substack{w^* \in \partial w(d) \\ \text{cfr. déf. } w(d)}} w^{*T}(d_{p\gamma} - d_\gamma) \\ &= \max_{u \in \partial\Delta^+(d_\gamma)} (\gamma u + d_\gamma)^T(d_{p\gamma} - d_\gamma), \end{aligned}$$

il existe un sous-gradient, disons $u^* \in \partial\Delta^+(d_\gamma)$, tel que

$$(\gamma u^* + d_\gamma)^T(d_{p\gamma} - d_\gamma) \geq 0. \quad (5.13)$$

De (5.11), tenant compte du fait que (5.12) tient pour $u = u^*$, on a :

$$\begin{aligned} & \gamma p\Delta^+(d_\gamma) + \frac{1}{2}\|d_\gamma\|^2 + (\gamma pu^* + d_\gamma)^T(d_{p\gamma} - d_\gamma) + \frac{1}{2}\|d_{p\gamma} - d_\gamma\|^2 + \gamma(1-p)\Delta^-(d_{p\gamma}) \\ & \leq \gamma p\Delta^+(d_{p\gamma}) + \frac{1}{2}\|d_{p\gamma}\|^2 + \gamma(1-p)\Delta^-(d_{p\gamma}) \\ & \leq \gamma p\Delta^+(d_\gamma) + \gamma(1-p)\Delta^-(d_\gamma) + \frac{1}{2}\|d_\gamma\|^2. \end{aligned}$$

On a donc :

$$(\gamma pu^* + d_\gamma)^T(d_{p\gamma} - d_\gamma) + \frac{1}{2}\|d_{p\gamma} - d_\gamma\|^2 \leq \gamma(1-p)[\Delta^-(d_\gamma) - \Delta^-(d_{p\gamma})],$$

qui peut être réécrit :

$$(\gamma u^* + d_\gamma)^T(d_{p\gamma} - d_\gamma) - (1-p)(\gamma u^* + d_\gamma)^T(d_{p\gamma} - d_\gamma) + (1-p)d_\gamma^T(d_{p\gamma} - d_\gamma) + \frac{1}{2}\|d_{p\gamma} - d_\gamma\|^2$$

$$\leq \gamma(1-p)[\Delta^-(d_\gamma) - \Delta^-(d_{p\gamma})].$$

Par conséquent, en tenant compte de (5.13), on a :

$$\begin{aligned} \|d_{p\gamma} - d_\gamma\|^2 &\leq 2(1-p)\{\gamma u^* + d_\gamma\}^T(d_{p\gamma} - d_\gamma) + \gamma[\Delta^-(d_\gamma) - \Delta^-(d_{p\gamma})] - d_\gamma^T(d_{p\gamma} - d_\gamma) \\ &\leq 4(1-p)(\gamma L + D), \end{aligned}$$

où L est la plus grande des constantes de Lipschitz entre Δ^+ et Δ^- ; D est le diamètre de l'ensemble où les points $d_{p\gamma}$ et d_γ sont situés ($d_{p\gamma}$ et d_γ sont bornés car γ_{max} est borné).

Finalement, puisque

$$|\Delta^+(d_{p\gamma}) - \Delta^+(d_\gamma)| \leq L\|d_{p\gamma} - d_\gamma\| \leq \sqrt{4(1-p)L(\gamma L + D)},$$

on obtient la thèse, à condition que $4(1-p)L(\gamma L + D) \leq \eta^2$, i.e. chaque fois que $p \geq 1 - \frac{\eta^2}{4L(\gamma L + D)} \triangleq p_{max}$. \square

Prouvons maintenant que l'itération principale se termine.

Lemme 5.4.2 *L'itération principale se termine après un nombre fini de pas.*

Preuve : Observons d'abord que l'algorithme ne peut pas boucler infiniment entre le pas 2 et 3. Ceci est une conséquence du lemme 5.4.1 et de la règle de mise à jour de p dans le cas $\Delta^+(d_{p\gamma}) - \Delta^+(d_\gamma) > \eta$ au pas 3.

Il ne nous reste donc plus qu'à montrer que l'algorithme ne peut pas passer une infinité de fois par le pas 4, ni par le pas 6, où respectivement un test de sortie et un test d'arrêt sont exécutés.

1. Commençons par montrer que l'algorithme ne peut pas passer une infinité de fois par le pas 4.

Supposons par l'absurde que l'algorithme passe une infinité de fois par le pas 4 et que la condition de descente (5.7) n'est jamais satisfaite. On passe donc aussi une infinité de fois par le pas 5, et indiquons par k toutes les quantités se rapportant au $k^{\text{ème}}$ passage.

Il est facile de vérifier que les règles d'insertion au pas 5 impliquent qu'il existe un indice \bar{k} tel que $\forall k \geq \bar{k}$, on assigne à tous les éléments

de faisceau nouvellement générés, un indice appartenant à I_+ (un indice appartenant à I_- est également possible). En effet, à force de passer dans le cas (a) du pas 5, on voit que comme γ diminue à chaque itération, ce dernier finira par passer en-dessous du seuil $\frac{\epsilon}{2\|g_{I_-}\|}$ et donc, par le lemme 5.2.1, on aura que $\|d_{p\gamma}\| \leq \epsilon$. On prend alors \bar{k} , le premier indice tel que l'on ait cela.

De plus, dans l'élément de faisceau inséré, on pose :

$$\hat{\alpha}_k = \max(0, \hat{\alpha}_k).$$

En outre, les règles d'insertion au pas 5 assurent aussi que la norme de $d_{p\gamma}$ est borné (car $\forall k \geq \bar{k}$, $\|d_{p\gamma}\| \leq \epsilon$), et par conséquent, par Bolzano-Weierstrass, il existe une sous-suite, disons $\{d_{p\gamma}^k\}_{k \in \mathcal{K}}$, qui converge vers une limite \hat{d} .

La sous-suite correspondante $\{\Delta^+(d_{p\gamma}^k)\}_{k \in \mathcal{K}}$ est bornée car $\Delta^+(d_{p\gamma}) \leq \max_{i \in I_+} \{g_i^T d_{p\gamma}\}$, or $d_{p\gamma}$ est borné et f est localement Lipschitzienne, ce qui entraîne que g_i est borné. La sous-suite $\{\Delta^+(d_{p\gamma}^k)\}_{k \in \mathcal{K}}$ étant bornée, elle admet donc à son tour une sous-suite convergente pour $k \in \mathcal{K}' \subseteq \mathcal{K}$, disons vers $\hat{\Delta}^+ \leq -\hat{\sigma}$, où $\hat{\sigma} \triangleq \lim_{k \rightarrow \infty} \sigma_k > 0$ ($\sigma_k = \frac{\gamma_{\min}^k \delta}{2}$ diminue car γ_{\min}^k diminue et est positif).

Soient maintenant s et t , deux indices successifs dans \mathcal{K}' . On a

$$\Delta^+(d_{p\gamma}^t) \geq \hat{g}^{(s)T} d_{p\gamma}^t - \max(0, \hat{\alpha}_s) \quad (5.14)$$

et

$$m\Delta^+(d_{p\gamma}^s) \leq \hat{g}^{(s)T} d_{p\gamma}^s - \max(0, \hat{\alpha}_s), \quad (5.15)$$

la dernière relation étant une conséquence de la définition de $\hat{\alpha}_s$ et de (5.8). En effet, dans le cas (5.8), c'est parce que $\max(0, \hat{\alpha}_s) = 0$ et dans les autres cas, c'est parce que

$$\begin{aligned} f(\hat{x}) > f(y) + m\Delta^+(d_{p\gamma}) &\iff f(\hat{x}) - f(y) > m\Delta^+(d_{p\gamma}) \\ &\stackrel{\text{déf. } \hat{\alpha}}{\iff} \hat{g}^T d_{p\gamma} - \hat{\alpha} > m\Delta^+(d_{p\gamma}). \end{aligned}$$

Les relations ci-dessus impliquent ((5.14)-(5.15)) en passant à la limite

$$(1 - m)\Delta^+ \geq 0,$$

ce qui contredit $\hat{\Delta}^+ \leq -\hat{\sigma} > 0$.

2. Afin de compléter la preuve, nous avons besoin de montrer que l'algorithme ne peut pas passer une infinité de fois par le pas 6. Supposons par l'absurde que l'algorithme passe une infinité de fois par le pas 6. On indice par $k \in \mathcal{K}$ toutes les quantités se rapportant au $k^{\text{ème}}$ passage.

On observe que $\forall k$, on a :

$$\Delta^+(d_\gamma^k) \geq \Delta^+(d_{p\gamma}^k) - \eta_k > -(\sigma_k + \eta_k) \quad \text{cfr. pas 3.} \quad (5.16)$$

De plus, chaque fois que le pas 6 est exécuté, γ_{max} est réduit selon (5.9). Donc, sachant que γ_{min}^k est décroissante monotone et bornée ($0 < \gamma_{min} < \frac{\epsilon}{2\|g_{I_-}\|}$), γ_{max}^k devient arbitrairement près de γ_{min}^k puisque l'on fait une infinité de fois le (5.9). Par conséquent, par le lemme 5.2.1, et comme $\gamma_{min}^k < \frac{\epsilon}{2\|g_{I_-}^k\|}$, on a asymptotiquement que $\|d_{p\gamma}^k\| \leq \epsilon$, qui à son tour entraîne que l'on assigne un indice appartenant à I_+ aux éléments de faisceau (un indice appartenant à I_- est également possible). En outre, dans l'élément de faisceau inséré, on pose :

$$\hat{\alpha}_k = \max(0, \hat{\alpha}_k).$$

On peut noter que par le critère de suppression du faisceau au pas 6 de l'itération principale, $a_i \leq \epsilon \quad \forall i \in I_+$.

Maintenant, de (5.16) et (5.4), on a :

$$\|d_\gamma^k\| \leq \sqrt{(\sigma_k + \eta_k)\gamma_k}$$

et

$$-\frac{1}{\gamma_k}d_\gamma^k \in \text{conv} \{g_i \mid i \in I_+^{(k)}\} \quad \text{par (5.4),}$$

ce qui implique

$$\|g_*^k\| \leq \sqrt{\frac{\sigma_k + \eta_k}{\gamma_k}} \leq \sqrt{\frac{\sigma_k + \eta_k}{\gamma_{min}^k}}.$$

Et ceci, sachant que $\delta \geq \sqrt{\frac{\sigma_k + \eta_k}{\gamma_{min}^k}}$ contredit le fait que l'algorithme ne se termine pas. \square

On va maintenant montrer que tout l'algorithme se termine.

Théorème 5.4.1 $\forall \epsilon > 0$ et $\forall \delta > 0$, l'algorithme s'arrête en un nombre fini d'itérations principales, en un point stationnaire satisfaisant la condition de stationarité

$$\|g_\star\| \leq \delta \quad \text{avec } g_\star \in \partial_\epsilon^G f(y).$$

Preuve : On doit montrer que la condition d'arrêt au pas 1 et au pas 6 est vérifiée en nombre fini d'exécutions de l'itération principale.

Supposons par l'absurde que ce n'est pas le cas. Par le lemme 5.4.2, il suit que la condition de descente (5.7) au pas 4 est satisfaite une infinité de fois. Soit $y^{(k)}$ le centre de stabilité à la $k^{\text{ème}}$ itération principale et indiquons par $k \in \mathcal{K}$ toutes les quantités se rapportant au $k^{\text{ème}}$ passage.

Puisque

$$f(y^{(k+1)}) \leq f(y^{(k)}) + m\Delta^+(d_{p\gamma}^{(k)})$$

.

.

.

$$f(y^{(k+1)}) - f(y^{(0)}) \leq m \sum_{i=0}^k \Delta^+(d_{p\gamma}^{(i)}),$$

et sachant que $\Delta^+(d_{p\gamma}^{(i)}) \leq -\sigma_k$ (cfr. pas 2) et que σ_k est borné loin de zéro, on obtient :

$$\lim_{k \rightarrow +\infty} f(y^{(k+1)}) - f(y^{(0)}) \leq -\infty,$$

ce qui est impossible car f est borné par le dessous par les hypothèses **H 1** et \mathcal{F}_0 compact. \square

5.5 Conclusions

On a décrit un modèle d'algorithme, pour la minimisation de fonctions non convexes et non différentiables, basé sur la construction simultanée de deux approximations affines par morceaux, et on a prouvé que l'algorithme converge vers un point stationnaire.

Cependant, pour avoir un algorithme implémentable pratiquement, plusieurs questions sont encore à traiter.

On mentionne en particulier le problème de trouver une technique d'agrégation de sous-gradients appropriée pour tenir compte de la capacité limitée de stockage ; le problème de sélection effective du paramètre γ à l'étape 2 de l'itération principale ; et finalement, celui de l'implémentation efficace de la recherche linéaire au pas 5.

Appendice : On discute le problème de trouver un minimum global au problème

$$\min_d f_{p\gamma}(d) \triangleq \min_d \gamma \Delta_p(d) + \frac{1}{2} \|d\|^2,$$

qui, grâce aux définitions données dans la section 2, peut être écrit :

$$\min_d \gamma [p \max_{i \in I_+} \{g_i^T d - \alpha_i\} + (1-p) \min_{i \in I_-} \{g_i^T d - \alpha_i\}] + \frac{1}{2} \|d\|^2.$$

Le problème ci-dessus consiste à minimiser une fonction objectif qui est la somme d'une fonction fortement convexe (non différentiable) et d'une fonction concave affine par morceaux.

On peut définir $\forall k \in I_-$ le problème d'optimisation convexe et non différentiable $P^{(k)}$ suivant :

$$\min_d f_{p\gamma}^k(d) \triangleq \min_d \gamma [p \max_{i \in I_+} \{g_i^T d - \alpha_i\} + (1-p) \min_{i \in I_-} \{g_i^T d - \alpha_i\}] + \frac{1}{2} \|d\|^2.$$

Il est évident que

$$f_{p\gamma}(d) \leq f_{p\gamma}^k(d) \quad \forall d \in \mathbb{R}^n \quad \text{et} \quad \forall k \in I_-,$$

et que $\forall d \in \mathbb{R}^n \exists k \in I_-$ tel que

$$f_{p\gamma}(d) = f_{p\gamma}^k(d).$$

D'où, il suit qu'un minimum global $d_{p\gamma}$ de $f_{p\gamma}(d)$ peut être localisé en trouvant le "meilleur" parmi tous les minima des fonctions $f_{p\gamma}^k(d)$, $k \in I_-$. Trouver le minimum global d'une fonction $f_{p\gamma}^k(d)$ demande à son tour de résoudre un problème quadratique convexe du type que l'on résout habituellement dans les méthodes de type faisceau.

En conclusion, la minimisation globale de $f_{p\gamma}(d)$ exige la solution de $|I_-|$ problèmes de programmation quadratique.

La condition (nécessaire et suffisante) pour que le point $d_{p\gamma}^{(k)}$ soit optimal pour le problème $P^{(k)}$ est :

$$0 \in \partial f_{p\gamma}^k(d_{p\gamma}^{(k)}), \text{ i.e. } 0 \in d_{p\gamma}^{(k)} + \gamma[p\partial\Delta^+(d_{p\gamma}^{(k)}) + (1-p)g_k]. \quad (5.17)$$

Donc, en prenant $K_- \subseteq I_-$ le plus petit des sous-ensembles d'indices telle que $d_{p\gamma} = d_{p\gamma}^{(k)}$, la condition (5.17) devient

$$0 \in d_{p\gamma} + \gamma[p\partial\Delta^+(d_{p\gamma}) + (1-p)g_k] \quad \forall k \in K_-,$$

qui peut être reconnue comme étant la condition d'optimalité locale pour les fonctions quasidifférentiables (voir [3]). Il est important de noter que, à cause de la structure particulière de la fonction $f_{p\gamma}(d)$, une telle condition est à la fois nécessaire et suffisante.

Chapitre 6

Résultats Numériques

La méthode faisceau proximale décrite dans le chapitre 4 a été implémentée en Matlab. Afin de montrer la fiabilité du code, quelques essais numériques ont été effectués sur un ensemble de 13 problèmes tests issus de [12]. Tous les calculs ont été fait sur un processeur AMD Athlon 2600+ avec 512 MB de mémoire RAM.

6.1 Problèmes tests

Présentons tout d'abord les problèmes tests issus de [12]. Afin de gagner de la place, on ne montrera que les solutions optimales des problèmes de dimension inférieure ou égale à 4, pour plus de détails voir par exemple Lemaréchal et Mifflin (1977), Shor (1985), Schramm (1989), Kiwiel (1990).

Rosenbrock (Rsbk)

Dimension	2,
Fonction objectif	$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$,
Point optimum	$x^* = (1, 1)$,
Valeur optimale	$f(x^*) = 0$,
Point de départ	$x^0 = (-1.2, 1)$.

Crescent (Crst)

Dimension	2,
Fonction objectif	$f(x) = \max\{x_1^2 + (x_2 - 1)^2 + x_2 - 1, -x_1^2 - (x_2 - 1)^2 + x_2 + 1\}$,
Point optimum	$x^* = (0, 0)$,
Valeur optimale	$f(x^*) = 0$,
Point de départ	$x^0 = (-1.5, 2)$.

CB2

Dimension	2,
Fonction objectif	$f(x) = \max\{x_1^2 + x_2^4, (2 - x_1)^2 + (2 - x_2)^2, 2e^{-x_1+x_2}\},$
Point optimum	$x^* = (1.139286, 0.899365),$
Valeur optimale	$f(x^*) = 1.952225,$
Point de départ	$x^0 = (1, -0.1).$

CB3

Dimension	2,
Fonction objectif	$f(x) = \max\{x_1^2 + x_2^4, (2 - x_1)^2 + (2 - x_2)^2, 2e^{-x_1+x_2}\},$
Point optimum	$x^* = (1, 1),$
Valeur optimale	$f(x^*) = 2,$
Point de départ	$x^0 = (2, 2).$

DEM

Dimension	2,
Fonction objectif	$f(x) = \max\{5x_1 + x_2, -5x_1 + x_2, x_1^2 + x_2^2 + 4x_2\},$
Point optimum	$x^* = (0, -3),$
Valeur optimale	$f(x^*) = -3,$
Point de départ	$x^0 = (1, 1).$

QL

Dimension	2,
Fonction objectif	$f(x) = \max_{1 \leq i \leq 3} f_i(x),$
où	$f_1(x) = x_1^2 + x_2^2,$ $f_2(x) = x_1^2 + x_2^2 + 10(-4x_1 - x_2 + 4),$ $f_3(x) = x_1^2 + x_2^2 + 10(-x_1 - 2x_2 + 6),$
Point optimum	$x^* = (1.2, 2.4),$
Valeur optimale	$f(x^*) = 7.2,$
Point de départ	$x^0 = (-1, 5).$

LQ

Dimension	2,
Fonction objectif	$f(x) = \max\{-x_1 - x_2, -x_1 - x_2 + (x_1^2 + x_2^2 - 1)\},$
Point optimum	$x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}),$
Valeur optimale	$f(x^*) = -\sqrt{2},$
Point de départ	$x^0 = (-0.5, -0.5).$

Mifflin1 (Mfl1)

Dimension	2,
Fonction objectif	$f(x) = -x_1 + 20 \max\{x_1^2 + x_2^2 - 1, 0\},$
Point optimum	$x^* = (1, 0),$
Valeur optimale	$f(x^*) = -1,$
Point de départ	$x^0 = (0.8, 0.6).$

Mifflin2 (Mff2)

Dimension	2,
Fonction objectif	$f(x) = -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75 x_1^2 + x_2^2 - 1 ,$
Point optimum	$x^* = (1, 0),$
Valeur optimale	$f(x^*) = -1,$
Point de départ	$x^0 = (-1, -1).$

Rosen

Dimension	4,
Fonction objectif	$f(x) = \{f_1(x), f_1(x) + 10f_2(x), f_1(x) + 10f_3(x), f_1(x) + 10f_4(x)\},$
où	$f_1(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4,$
	$f_2(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8,$
	$f_3(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10,$
	$f_4(x) = x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5,$
Point optimum	$x^* = (0, 1, 2, -1),$
Valeur optimale	$f(x^*) = -44,$
Point de départ	$x^0 = (0, 0, 0, 0).$

Maxquad (Maxd)

Dimension	10,
Fonction objectif	$f(x) = \max_{1 \leq i \leq 5} \{x^T A_i x - b_i^T x\},$
Valeur optimale	$f(x^*) = -0.8414084,$
Point de départ	$x^0 = (1, \dots, 1).$

Maxq

Dimension	20,
Fonction objectif	$f(x) = \max_{1 \leq i \leq 20} x_i^2,$
Valeur optimale	$f(x^*) = 0,$
Point de départ	$x_i^0 = i/10, \quad i = 1, \dots, 10,$ $x_i^0 = -i/10, \quad i = 11, \dots, 20.$

Maxl

Dimension	20,
Fonction objectif	$f(x) = \max_{1 \leq i \leq 20} x_i ,$
Valeur optimale	$f(x^*) = 0,$
Point de départ	$x_i^0 = i/10, \quad i = 1, \dots, 10,$ $x_i^0 = -i/10, \quad i = 11, \dots, 20.$

6.2 Résultats numériques

Les résultats obtenus sont repris dans le tableau 6.2.1.
On utilisera les abréviations suivantes :

Pbme	-	le nom du problème test
it	-	le nombre d'itérations ou pas sérieux
x^*	-	la valeur finale du centre de stabilité
f^*	-	la valeur finale de la fonction objectif
t	-	le temps d'exécution du programme.

Les paramètres d'entrée ont été initialisés comme suit : $\epsilon = 0.1$, $\delta = 10^{-5}$, $m = 0.2$, $\rho = 0.5$, $r = 0.5$, $R = 10^3$. En testant l'algorithme, on a toujours gardé le même ensemble de paramètres d'entrée, sans changements basés sur un problème test spécifique, dans le but de contrôler la robustesse plus que l'efficacité.

Quelques changements majeurs ont été effectués dans le code par rapport à l'algorithme 4.3.1.

Le premier concerne la condition d'arrêt au pas 3. En effet, le calcul de g^* s'est avéré demander un temps de calcul beaucoup trop élevé et a donc été remplacé par une autre condition, portant sur l'écart entre deux centres de stabilité consécutifs.

Le second changement se produit lorsqu'il y a beaucoup de pas nuls, car il peut alors arriver que le $\hat{\gamma}$ diminue au point d'être égal à γ_{min} , entraînant de ce fait un bouclage à l'infini du programme (le $\hat{\gamma}$ restant inchangé, la solution du problème $QP(\hat{\gamma})$ reste inchangée également). L'issue trouvée à cela, a été de diminuer le γ_{min} .

Tableau 6.2.1

Pbme	it	x^*	f^*	t (sec)
Rsbk	790	(1.000055, 1.000110)	0.000000	52.385000
Crst	21	(0.000333, -0.001106)	0.001107	12.238000
CB2	17	(1.139039, 0.899559)	1.952224	1.332000
CB3	11	(1.012800, 0.980951)	2.014458	18.457000
DEM	12	(-0.000000, -3.000000)	-3.000000	0.500000
QL	54	(1.199974, 2.40001)	7.200000	1.78300
LQ	8	(0.707107, 0.707107)	-1.414214	0.471000
Mfl1	42	(1.000000, 0.000031)	-1.000000	2.574000
Mfl2	38	(1.000000, 0.000023)	-1.000000	1.693000
Rosen	20	(-0.000006, 0.999993, 2.000006, -0.999994)	-44.000000	1.812000
Maxd	43	(-0.126258, -0.034379, -0.006857, 0.026363, 0.067295, -0.278397, 0.074219, 0.138523, 0.084030, 0.038579)	-0.841408	175.632000
Maxq	169	(-0.000078, -0.000077, -0.000014, -0.000040, -0.000070, 0.000002, 0.000103, 0.000058, -0.000106, 0.000068, 0.000127, 0.000027, -0.000157, -0.000196, -0.000024, -0.000036, -0.000024, -0.000119, 0.000037, -0.000024)	0.000000	22.352000
Maxl	212	(-0.000000, -0.000098, -0.000098, -0.000000, 0.000098, 0.000000, 0.000098, 0.000098, 0.000098, 0.000098, 0.000000, -0.000098, 0.000000, 0.000000, -0.000000, 0.000000, 0.000000, -0.000098, 0.000000, 0.000000)	0.000000	10.205000

Bibliographie

- [1] F. Clarke. Optimization and nonsmooth analysis. John Wiley and Sons, 1983.
- [2] M. Gaudioso et A. Astorino et A. Fuduli. Analysis of regularization techniques in convex nondifferentiable optimization. Operations Research Proceedings 1996, Springer-Verlag, Berlin, Heidelberg, pp. 20-25.
- [3] V.F. Demyanov et A. Rubinov. Constructive nonsmooth analysis. Verlag Peter Lang, 1995.
- [4] E.W. Cheney et A.A. Goldstein. Newton's method for convex programming and Tchebycheff approximation. Numerische Mathematik, 85, pp. 313-334, 1959.
- [5] J.-P. Hiriart-Urruty et C. Lemaréchal. Convex Analysis and Minimization Algorithms, Vol. I-II. Springer-Verlag, Berlin, Heidelberg, New-York, 1993.
- [6] E. Polak et D. Mayne et J. Higgins. A superlinear convergent algorithm for min-max problems. Proceedings of the 28th Conference on decision and control, Tampa, Florida, December 1989, pp. 894-898.
- [7] H. Schramm et J. Zowe. A version of the bundle idea for minimizing a nonsmooth function : conceptual idea, convergence analysis. Siam Journal on Optimization, 1, pp. 121-152, 1992.
- [8] G.D. Pillo et L. Grippo et S. Lucidi. A smooth method for the finite minimax problem. Mathematical Programming, 60, pp. 187-214, 1993.
- [9] A. Fuduli et M. Gaudioso. The proximal trajectory algorithm for convex minimization. Submitted for publication, 2001.
- [10] A. Fuduli et M. Gaudioso et G. Giallombardo. Minimizing nonconvex nonsmooth functions via cutting planes and proximity controls. Submitted for publication, 2002.
- [11] A. Fuduli et M. Gaudioso et G. Giallombardo. A DC piecewise affine model and a bundling technique in nonconvex nonsmooth minimization. Submitted for publication, 2003.

- [12] Marko M. Mäkelä et Pekka Neittaanmäki. Nonsmooth Optimisation. World Scientific, 1992.
- [13] J.W. Daniel et W.B. Gragg et L. Kaufman et G.W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation*, 30 (136), pp. 772-795, 1976.
- [14] A. Frangioni. Solving semidefinite quadratic problems within nonsmooth optimization algorithms. *Computers and Operations Research*, 23 (11), pp. 1099-1118, 1996.
- [15] A. Fuduli. Metodi numericiper la minimizzazione di funzioni convesse nondifferenziabili. Ph. D. Thesis, 1998.
- [16] M. Gaudioso. Nonsmooth Optimization. Handbook of applied optimization, M. G. C. Resende and P. Pardalos, eds., Oxford University Press, New York, pp. 299-310, 2002.
- [17] J.E. Kelley. The cutting plane method for solving convex programs. *Journal of the SIAM*, 8, pp. 703-712, 1960.
- [18] K.C. Kiwiel. An aggregated subgradient method for nonsmooth convex minimization. *Mathematical Programming*, 27, pp. 320-341, 1983.
- [19] K.C. Kiwiel. Methods of descent for nondifferentiable optimization. Springer-Verlag, Berlin, Heidelberg, 1985.
- [20] K.C. Kiwiel. Finding normal solutions in piecewise linear programming. *Applied Mathematics and Optimization*, 32, pp. 105-122, 1995.
- [21] K.C. Kiwiel. Restricted step and Levenberg-Marquardt techniques in proximal bundle methods for nonconvex nondifferentiable optimization. *Siam Journal on Optimization*, 6, pp. 227-249, 1996.
- [22] K.C. Kiwiel. A bundle Bregman proximal method for convex nondifferentiable optimization. *Mathematical Programming*, 85, pp. 241-258, 1999.
- [23] C. Lemaréchal. A view of line-searches. Optimization and optimal control, W. O. A. Auslender and J. Stoer, eds., vol. 30 of Lecture notes in control and information sciences, Springer-Verlag, pp. 59-78, 1981.
- [24] R. Mifflin. An algorithm for constrained optimization with semi-smooth functions. *Mathematics of Operations Research*, 2, pp. 191-207, 1977.
- [25] M. Mäkelä. Survey of bundle methods for nonsmooth optimization. *Optimization Methods and Software*, 17, pp. 1-29, 2002.
- [26] R.T. Rockafellar. Convex Analysis. Princeton University Press, 1970.
- [27] N. Shor. Minimizations Methods for Nondifferentiable Functions. Springer-Verlag, Berlin, 1985.

Tout d'abord, je voudrais remercier mon promoteur, Mr J.J. Strodilot, pour l'aide et le temps qu'il a bien voulu m'accorder tout au long de l'année. Il a toujours tenu à ce que je l'informe de l'avancée de ce mémoire et m'a secouru quand besoin était. Pour cela, je l'en remercie.

Je remercie ensuite ma famille, pour son soutien psychologique non seulement au cours de cette année, mais pour toutes mes années d'études, qui n'ont pas été de tout repos.

Je remercie enfin mes amis et camarades de classe d'avoir toujours su garder une bonne ambiance et d'avoir ainsi rendu ces quatre années agréables.